DOCUMENT RESUME

ED 089 361 CS 201 191

AUTHOR Dale, Edgar, Ed.
TITLE Readability.
INSTITUTION National Conference on Research in English.; National Council of Teachers of English, Champaign, Ill.
PUB DATE 49
NOTE 46p.

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS Elementary Education; *Evaluation Criteria; *Literature Reviews; *Readability; Reading Difficulty; *Reading Instruction; Reading Material Selection; *Reading Research
IDENTIFIERS Typography

ABSTRACT

This monograph discusses the three chief interacting variables which affect the readability of a particular piece of material--format and organization, the reader, and the criterion used to estimate readability. Chapters include "The Concept of Readability," which provides a comprehensive view of readability, analyzes the purposes of research in readability, and discusses the strengths and weaknesses of present knowledge; "Readability Formulas--An Evaluation," which examines the criteria for reading difficulty and comprehension levels, citing some of the short-comings of current evaluation techniques; "The Use of Vocabulary Lists in Predicting Readability and in Developing Reading Materials," which considers the most common words used at differing reading levels, analyzes several vocabulary lists, and discusses the preparation of materials for certain levels of readability; "Typography and Readability," which examines the effect of type face, spatial arrangement of page, and other variables in format; and "Techniques for Selecting and Writing Readable Materials," which shows how teachers, librarians, and community leaders can use readability research to select reading materials for specific readers. Each chapter contains a bibliography. (RB)

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED 089361

Readability

Edited by Edgar Dale

A publication of the
National Conference on Research in English

Reprinted from
ELEMENTARY ENGLISH
(January to May, 1949).
An official publication of
the National Council of Teachers of English,
704 S. Sixth Street, Champaign, Ill.
John J. DeBoer, Editor

60c Per Copy

50c Each For 10 or More

AS 201 191

TABLE OF CONTENTS

Page

The Concept of Readability 1
EDGAR DALE AND JEANNE S. CHALL

Readability Formulae - An Evaluation 8
IRVING LORGE

The Use of Vocabulary Lists 17
E. W. DOLCH

Typography and Readability 26
HAROLD E. BURTT

Techniques for Selecting and Writing Readable Materials 36
EDGAR DALE AND JEANNE S. CHALL

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY

National Council of Teachers of English

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER.

The Concept of Readability

EDGAR DALE AND JEANNE S. CHALL[1]

The impressive percentages of radio owners, the amount of time they spend listening to the radio, and looking at television, and the figures on movie attendance sometimes obscure the fact that adults still read. In fact, a recent survey of the leisure-time habits of adults in 17 cities disclosed that reading is the most popular single pastime among city people in the United States. (8)*

The predictions, then, that radio and motion pictures would replace the printed page as the major medium of communication have not materialized. While it is true that radio, recordings, motion pictures, and filmstrips are being brought into the schools at an increasing rate, reading continues to be the major teaching tool of our schools.

Since reading is still the chief means whereby persons gain information, skills, and entertainment, the effectiveness with which books, newspapers, magazines, and pamphlets convey this information remains an important problem. Newspapers and magazines are increasingly aware of their deficiencies as instruments for communicating ideas and have recently shown concern in developing readable stories. The early attempts to make their publications readable by intuitive and rule-of-thumb methods have given way to an acceptance of the objective methods of measuring readablity—methods that have arisen through an accumulation of research carried on within the past twenty years.

This accumulated research has come primarily from educators who from the earliest days needed some means of selecting books for the different grades. That the first grade reader must be simpler than the second grade reader was self-evident, but the method of discovering proper gradients for books was not easily determined.

Research in readability has come also from leaders in adult education and from librarians who tried to find some means of "putting the right book into the hands of the right reader." Since adult education is often that of relating a man to a book —the book must fit the man if any education is to take place. In other words, the book must be readable.

What is a Readable Book?

But what do we mean by readability or a readable book? Webster's Unabridged Dictionary defines *readable* as: "legible," "easy to read because interesting or pleasing," "that permits or admits of reading." Obviously, there is much room here for confusion. Thus readablity may refer to legibility, interest, or ease of reading, or some combination of these. That confusion does exist can be seen from the blanket statements of book reviewers who very often state that the book being reviewed is "readable." Do they refer to its legibility, to its possible interest and ease? And

*Numbers in parentheses refer to references in Bibliography, pp. 25-26.

[1]Members of the faculty of the College of Education, Ohio State University. This is the first of a Series of five articles on readability sponsored by the National Conference on Research in English.

if they refer to the latter, we can also ask for whom is it interesting or easy? We don't all find the same books interesting and easy going.

To clear up some of the confusion and to get some common referents for the notion of readability, Gray and Leary asked groups of librarians, publishers, and teachers what in their opinion makes a book readable. (5) They received hundreds of descriptive statements but in general these judges agreed that factors of content were most important: factors of style next, format was third, and organization was last.

Bernice Leary, in a subsequent summary of the findings of this survey made these comments:

> ...According to the combined opinion of these judges, then, if you give a reader a theme that interests him, whether it be serious or trifling, whether it concern people, travel, adventure, science, or business, you have made a strong attack upon the problem of readability. If in addition, you discover what style of expression is best suited to the reader's needs and tastes, that is, the scope of vocabulary and the kind of sentences which he reads easily, and the type of approach that pleases him, you have the final solution of the problem close at hand. In the opinion of these judges the attractiveness of the book, its mechanical set-up, and its general plan of organization are matters of minor importance. (7, p. 280).

The opinions of adult readers in libraries as to what makes a book readable agreed closely with those of the librarians, teachers, and publishers. There was a difference, however, in what was thought most important. The readers considered style of expression most important, content second. A later survey of the opinions of high school students by Ruth Strang also disclosed stylistic factors as of first importance. (11)

The findings of Gray and Leary point to at least three broad aspects of readability—the content or subject matter and its interest and appeal to the reader: the style of expression which makes it comprehensible and interesting to him: the format and organization which make it easy to follow the logic of the material with a minimum of effort.

Typographical Aspects of Readability.

Research workers in readability concentrated their efforts within one or another of these broad aspects of readability. The earliest research was in connection with typography, which falls within the format category of Gray and Leary. The research in typography started with an interest in the relative legibility, from a distance, of isolated letters in different faces of printing type. It has since considered the problems of legibility at a natural reading distance. Tinker and Paterson in their book, *What Makes Type Readable*, define their concept of readability as follows: "In general, we have used the words legibility and readability interchangeably to mean 'ease and speed of reading printed material at a natural reading distance.' " (9, p. xvii)

Their criterion of readability or legibility has been the speed with which people can read printed matter. They have also used readers' judgments of legibility as a criterion. Readers' judgments coincided fairly well with the objective speed of reading measures. However, in a later study, these investigators found very close agreement between judged legibility and judged "pleasingness." (16) In general, the researchers in this aspect of readability are pretty much in agreement as to style of

type, the spacing between lines, the length of lines, the size of type, the width of margins, color of paper, etc., which make for optimal speed and ease of reading.

Although these researchers have recognized the influence of subject matter and expressional elements on speed and ease of reading, they have kept these factors constant while they studied only the typographical variables.*

Interest and Readability.

From an historical viewpoint, the next broad aspect of readability to receive attention was that of content and its effect upon interest. Teachers and librarians have long compiled lists of books for different age and grade levels. These lists were usually based upon the observed preferences of children in the particular age or grade groups.

The early construction of "interest" lists finally lead to a more detailed analysis of the elements within the books that make for interest or lack of interest. Dunn (2) and Gates (4) and others contributed considerably to our knowledge of what makes certain books interesting or uninteresting to children in the primary grades. Zeller (14) did a similar analysis for the junior high school level; Sterner (10) for the senior high school level; Gray and Munroe (6) and Waples and Tyler (13) for adults.

These researchers answered some questions as to the type of subject matter, the themes, and the elements in reading materials that appeal to different groups of readers. For example, Gates found that such elements as "surprise" (unexpectedness, unforseen events), "liveliness," (action, movement), "animalness" (presentations of things animals do) contribute positively to young children's interest. On the other hand the element of "moralness" decreases interest. (4, pp. 89-90) Zeller found the two factors of "action," and "humor" to exert the greatest influence on the reading choices of junior high school students. (14, p. 73) Sterner found recently that among high school students "adventure is the favorite with adolescents; humor is a close second, and the love theme is very popular with high school girls . . . ." (10, p. 60) Waples and Tyler list the subjects that different groups of adults prefer. (13)

But interest or appeal does not depend solely upon the subject matter and theme. It depends also upon the format since the style and size of type, the size of the book, its illustrations, etc. do have some effect upon the reader's choice and enjoyment of the book. Interest depends also upon the stylistic or expressional elements in the book. In fact, one of Gates' factors of interest was "difficulty." (4, p. 90) In other words, a given book may lose its interest for young children even though it has the elements of "surprise," "liveliness," etc., if the presentation has complicated sentence structure, abstract conceptions, and too great concentration of ideas.

That comprehensibility influences and often limits readers' preferences and actual choice of books has been pointed out by other investigators, particularly by Waples and Tyler.

Readability and Style.

This leads us to our third aspect of readability—factors of style of expression. Workers in this area sought answers to

*A comprehensive treatment of typography and readability appears on pp. 26-35 of this pamphlet.

such questions as "What kind of vocabulary, sentence structure and other expressional elements best suit the abilities of particular groups of readers?" For these workers the concept of readability is largely one of comprehensibility. That is, a book is readable for a particular group when it is comprehensible or understandable to that group. The group may be a first grade class, a 7th grade class in general science, or the reading public of a particular newspaper.

The criterion of readability in the sense of comprehensibility is some measure of understanding what is read. This has been determined in several ways. Vogel and Washburne (12) used the average reading ability of children who had read and liked a particular book as the difficulty index of that book. Dale and Tyler (1), Gray and Leary (5), and others tested a variety of passages on adult readers and assigned to each passage the average reading ability of those adults who had succeeded in answering correctly a specific number of questions on each passage.

After the index of difficulty of each passage is determined, the internal elements in each passage (the vocabulary, sentence structure, etc.) are correlated with the criterion, to determine which ones contribute to ease or difficulty. This procedure has resulted in numerous "readabilty formulae" or regression equations which can be used to predict the comprehensibility of written material by counting and weighting certain significant structural elements in the text. The prediction is often given in terms of grade levels or grade placements. For example, a formula may "place" a book at the 6th grade. This means that average 6th grade readers should be able to read the book with adequate understanding.

The concept of readability as comprehensibility has received the most emphasis recently. And usually, when the term readability is used today, it is this particular concept which is meant. We must remember, however, that the readability formulae which are based upon this concept of readability do not tell the entire story of comprehensibility. Most of the formulae measure comprehensibility by some measure of vocabulary load and sentence structure. Some use a measure of the relative number of ideas and of human interest. However, none of them adequately account for conceptual difficulty, semantic variations of commonly used words, etc.†

Another obvious shortcoming of viewing readability as comprehensibility alone is the fact that the available measures of comprehensibility do not consider the appeal of subject matter. It is a common observation among teachers and layman that when there is a strong interest in the subject matter, more effort is put forth and there is a greater amount of understanding. Measures of comprehensibility also neglect to account for the format and organization of materials. Note how easily we give up reading a passage which goes on for a page or more with no paragraphing.

Comprehensive View of Readability

We see, therefore, that although research workers in readability have emphasized one of the three broad aspects of

†See subsequent article by Dr. Irving Lorge on the history and evaluation of readability formulae. See also article by Dr. E. W. Dolch on the use of word lists in predicting readability, p. 17.

readability outlined above, these aspects are not mutually exclusive. They are all interrelated and have been seperated only for purposes of discovering the factors that result in the "success" that people have with a particular piece of written material.

In the broadest sense, then, readability is the sum total (including the interactions) of all those elements within a given piece of printed material that affects the success that a group of readers have with it. The success is the extent to which they understand it, read it at an optimum speed, and find it interesting.

Now, success depends upon other things besides the printed material itself. It depends upon the reader—his skill in reading, his intelligence, his experience, his maturity, his interest and purpose in reading. A book that is readable for a skilled reader is unreadable for one less skilled.

We can also have the situation where two readers may be equally skilled in reading as judged by a standardized reading test, and yet a book on a specialized subject, such as nuclear physics, may not be equally readable for both. The one with more background information in nuclear physics and with a special interest in it will find the book more readable. A sixteen year old farm boy might read a pamphlet on how to raise potatoes with interest, ease, and understanding. An equally able city boy might find the pamphlet hard and uninteresting.

Readability also depends upon the kind and degree of success that we wish a group of readers to have with the material. For example, if we are concerned only that a book be comprehensible for a particular grade, we must still ask: How comprehensible should it be? When we select a book for the fourth grade, should it be comprehensible to every child in that grade?

Obviously, then, we shall have to select a book that will be below the ability of almost all of the children in that grade. If we decide that it should be comprehensible for the majority, then the book selected will be more difficult than if we had chosen one comprehensible for all. In short, then, a measure of comprehensibility is a relative one. The same can be said of interest.

Summary

We have discussed the three chief interacting variables which affect the readability of a particular piece of material. First, the book or article itself—its format and organization; its subject matter and themes; its expressional elements such as vocabulary, sentence structure, etc. Second, the reader—his general reading ability, his interest and purpose in reading, his general experience and specific experience along the lines of the book he is reading Third, the criterion used to estimate readability—whether we use a measure of interest, comprehension or speed of reading; and the methods used to estimate these criteria.

Applications of Research in Readability

What is the purpose of all this? Why all the time and effort spent in an area of research that has given us only approximations? The purpose, as in all scientific investigation, is prediction and control. Although we have not achieved the tools whereby to predict and control an individual's success with a particular book, we do have some tools, although still rough, to

predict and control the success that certain groups will have with particular books, articles, etc.—on the basis of interest, comprehension, and speed.

For the school, this is particularly important at all grades. Beginners in reading must have success with their books. Difficult, uninteresting, and unattractive books may have a detrimental effect upon their learning to read. As a result of the work of Gates and his students, the rate of adding new words in primers has greatly changed. The work of Dunn and Gates on children's reading interests has also contributed to the production of books within the interests and needs of young children.

Once the child has mastered the basic skills of reading and can use reading as a tool for learning specific subject matter, the importance of having readable books is equally great. Selecting readable books for a class in the upper elementary grades is harder than selecting books for the lower grades. This is because of the wider range of abilities within one particular class as we go up the scale. Within a 6th grade class, for example, the reading abilities may range from the 3rd grade to over 11th grade. Besides this rather obvious difference in reading ability, there is also an increasing difference arising from experience, interests, etc. Research in readability has emphasized these differences and has contributed to the accepted notion that more than one book be used so that all children can work within their abilities and interests.

Readability research has contributed extensively to adult education. First, it has provided the techniques for the selection and preparation of materials for literacy and citizenship classes. The same principles of controlled vocabulary, simple sentence structure, etc. have been used to assure optimum success in learning to read. The findings of the studies on adult reading interests have also been helpful in preparing and selecting simple materials that will appeal to the mature interests of the readers. The findings will also have value to those interested in teaching English as a foreign language.

Studies in readability have also brought to the attention of educators and publishers the lack of simply written books on serious subjects that are of interest to adults with limited reading ability. In fact, the work of Bryson (15), the studies of Gray and Leary, Dale and Tyler, and Flesch (3) were specifically undertaken to determine how such materials can be better written.

We often criticize the reading habits and tastes of certain groups in the adult population—the readers of true story magazines, mystery magazines, and comic books. Yet, we have failed to provide an abundance of serious, interesting, easy reading materials for them.

Readability research is influencing the style of newspaper and magazine writing, and the publications of government and private educational agencies. Their early interest in legibility and human interest has recently given way to an interest in the comprehensibility of their publications. Rudolf Flesch's current work with the Associated Press shows the pervading influence of a body of research that began with problems of selecting textbooks for children.

Strengths and Weaknesses of Our Present Knowledge

All the research and application in the field of readability has gained much publicity and acceptance. It is no longer necessary to explain exactly what we mean by readability when talking to teachers and newspaper people. But, do we have all the answers? Have we considered enough factors when we make a statement that book X is more readable than book Y or that newspaper Z is the most readable in the country? We must admit that we have not. But we have made some inroads into the problem. For example, we have some tested generalizations as to typography which results in optimum speed and ease of reading.

On the problem of interest, we are still subject to the individual whims and tastes of the reader. However, we can profitably use the studies on interests and these findings can give us a clue as to whether we are at least satisfying a large proportion of a particular group of readers.

On the problem of comprehensibility, we have made some strides. The numerous readability formulae help give a rough approximation of the difficulty of a piece of material. Some of these are very easy to apply—and consist of mechanical counting of words, syllables, length of sentences, prepositional phrases, etc. But because they are mechanical, they are usually taken as infallible. The important factors of conceptual difficulty, organization of the material or the logic, semantic variations in words, etc. have been discussed widely in the literature in readability, but have not yet been incorporated in any formula.

To-date, there is no composite method that can be used to measure all aspects of readability. We must consider separately the aspects of format and organization, content, expressional elements; and then make a judgment as to the suitability of a particular book for a particular group.

Some day, though, we may be able to say that a given piece of material is readable for a particular group of readers and have this statement encompass all the possible factors that contribute to its readability. At the present time, however, we can say only that it is readable on the basis of such and such a criterion taking such and such factors into consideration.

BIBLIOGRAPHY

1. Dale, Edgar and Tyler, Ralph W. "A Study of the Factors Influencing the Difficulty of Reading Materials for Adults of Limited Reading Ability," *Library Quarterly*. IV (July, 1934), pp. 384-412.
2. Dunn, Fannie W. *Interest Factors in Primary Reading Material*, Contributions to Education, No. 113. New York: Bureau of Publications, Teachers College, Columbia University, 1921.
3. Flesch, Rudolf, *Marks of Readable Style*, Contributions to Education, No. 897 New York: Bureau of Publications, Teachers College, Columbia University, 1943. See also *Art of Plain Talk*, New York: Harper Bros., 1946, and "A New Readability Yardstick," *Journal of Applied Psychology*, XXXII (June 1948) pp. 221-233.
4. Gates, Arthur I. *Interest and Ability in Reading*. New York: The Macmillan Company, 1930.
5. Gray, William S. and Leary, Bernice E. *What Makes a Book Readable*. Chicago: University of Chicago Press, 1935.
6. Gray, William S. and Munroe, Ruth. *The Reading Interests and Habits of Adults*. New York. Macmillan Co., 1929.
7. Leary, Bernice E. "Difficulties in Reading Material" *Reading in General Education*. Washington, D. C.: American Council on Education, 1940, pp. 272-306.
8. *Opinion News*, December 1, 1947.

(Continued on Page 25)

Readability Formulae - An Evaluation

IRVING LORGE[1]

Teachers have ever been concerned about how well ideas get over in the classroom. To estimate their own success, they have used not only tests but also hunches. They have judged their own adequacy most frequently by their interpretation of children's performances on written examinations. Of course, many teachers supplement such appraisals by inferences based on the gleam in a pupil's eye, the attention in a classroom or the provocative discussion in an afternoon forum.

Teachers have been more critical, however, of the intended communication of others. They have tried to assess the potential success of the chart, the film, or the textbook, particularly the textbook. Uppermost was the question "Can our pupils understand this?" How often have teachers criticized a textbook with the complaints, "This is over their heads." "They can't get it." "What do they think the children are?"

Since the turn of the century, more and more attention has been devoted to obtaining valid, reliable, and impartial ratings of textbooks. In a large measure the attempts to appraise textbooks gave rise to the modern interest in readability. Vogel and Washburne's basic research in the estimating of the grade placement of children's reading materials established not only the fundamental concept of readability but provided the general method of measuring it. (1)

The concept of readability involves the idea of understanding printed material. For Vogel and Washburne, the idea implicit in the readability index of a text is the average amount of reading ability needed to understand the text. Specifically, Washburne and Vogel (2) attempted to classify books for appropriate grades. Instead of using the more conventional method of judging the relative difficulty, they made an empirical classification based not only on the expressed wishes of children but also on the measured reading ability of the children making the choices.

About thirty seven thousand children, each of whom had taken the paragraph-meaning section of the *Stanford Achievement Test,* were asked to fill out a ballot about the books they had read during the school year. Each child indicated each book he had liked in his year's reading. For some seven hundred different books at least twenty-five children had indicated they had read the book and liked it, too. The authors of the *Winnetka List* assumed that an average of the paragraph-meaning test scores would be indicative of the relative readability of the book. The *Winnetka List,* therefore, gives for each of seven hundred books the grade equivalent for the median paragraph meaning score. The *Winnetka Graded Book List* puts into relative order the books that children read and enjoy in the grade range from II to XI.

[1]Dr. Lorge is professor of education at Teachers College, Columbia University.

The grade placement of the book, therefore, represents the average reading ability of those children who read the book and who enjoyed it. The fact that the book was read implies that the book was available, that it was selected to satisfy some interest or need of the child, and that the book did satisfy the child enough for him to say he liked it. Since the grade placement is based on an average, it neglects the variation in the measurable reading ability of those who read it. Dull children may have read it and enjoyed it for its illustrations; brighter children may have read it and enjoyed it for the delineation of character or for the information it yielded. The average reading ability, therefore, is far from an uncontaminated measure of reading difficulty. By and large, however, it is a useful means of putting books together in appropriate age or grade classifications.

But what of the books that had not been included in the seven hundred? Was there an empirical and objective method for classifying the new books? The contribution of Vogel and Washburne was to relate their grade placement index to some features or characteristics of the book itself. Since the publication of E. L. Thorndike's *Teachers' Word Book* of 1921 (3), it was known that the more difficult books had more uncommon words. Washburne and Vogel sought for factors other than commonness of vocabulary that would differentiate among books of varying difficulty. They investigated such factors as the relative number of different words in the books, the kinds of sentences used, the relative number of prepositions and other elements of structure within the book.

They found that the greater the number of different words per thousand words of text, the higher the grade placement index. Longer sentences and an increase in the relative frequency of prepositions also made for a higher score. If per thousand running words of text, there were a large number of different words, a large number of uncommon words, a large number of prepositions, and many long sentences, the passage obtained a high grade-index; if the passage was limited to common words, few prepositions, and short and simple sentences, the grade-index was low.

Vogel and Washburne put these facts together in an empirical formula (statistically, a multiple regression formula) so that if one knew the relative number of each of the constituents, a reasonable *estimate* of grade placement would be made. Their 1928 formula estimated the grade index on the basis of just four factors:

1) the number of different words per thousand words of text
2) the number of uncommon words per thousand
3) the number of simple sentences in seventy-five successive sentences
4) the number of prepositions per thousand words

The utility of statistically-derived empirical formulae requires an understanding of their assumptions and limitations. Fundamentally, the readability formula can only be applied to books like those evaluated in the course of its development. Since the books considered in the Winnetka formula were those in children's popular reading, the formula can be used, most correctly, only for the estimation of the grade placement of children's reading. The more remote from children's reading, the less

adequate does the prediction become. The grade placement index for *Knickerbocker's History of New York*, of *Alice's Adventures in Wonderland* would probably be reasonable, but those for Hogben's *The Nature of Living Matter*, Mumford's *Technics and Civilization* or Dewey's *Logic* would be far from reasonable.

The difficulty in understanding books such as those by Hogben or Mumford or Dewey does not depend only on such elements of the writer's skill as his choice of vocabulary or his sentence structure. The comprehensibility of much of adult nonfiction is related more to characteristics of the reader and the quality of the ideas expressed in the book. The interpretation of the expressed thought is related more to the reader's informational background and motivations than to the internal evidences of the expressional facility of the author. The uncritical application of a readability formula to mature writing may reveal little but diligence in the applier.

The ever-present danger in such diligence is that the user of readability formulae may consider the number called the grade placement score as meaning measured reading difficulty. Grade placement should not neglect the maturity of the reader: his interests, his concept mastery, his drive, etc. Maturity of reading interest, moreover, cannot be considered the same as reading difficulty. The risk in the semantic confusion of "grade placement" with "reading difficulty" is always greater for books intended for adult levels than it is for books for children.

Vogel and Washburne, of course never implied that their formula could be applied to any book, or that it provided an over-all appraisal of reading difficulty. They distinguished clearly between "structural difficulty" and "content difficulty." Vogel and Washburne consider their notion of *structural difficulty* as appraised best by evidences of *expressional difficulties* within the text. Such expressional barriers to understanding are usually revealed by a relatively large number of different words, of uncommon or even rare words, of complicated sentence structure, and of a large proportion of prepositional phrases. Content difficulty is more difficult to define and even much more difficult to measure. It has to do with the maturity of the concepts, the recency of the experiences written about, and the cultural community of communication. Washburne and Vogel indicate that their formula is not concerned with "content difficulty" nor with the difficulty of the concepts, ideas, and generalizations. In the concluding paragraphs of their 1928 article they state:

"Any book for use in the *elementary grades* may be similarly analyzed. It is, therefore, possible to determine the correct grade placement for any book so far as *structural difficulty* is concerned." (p. 380 italics mine)

Essentially, then, the Vogel and Washburne Grade Placement Formula tried to predict a criterion, e. g. some aspect of difficulty on the basis of observable variables, e. g. kind of sentence, etc. The pattern established in this formula has been followed by Lewerenz (4) in 1929 and later, by Ojemann (5) in 1933, by Dale and Tyler (6) in 1934, by Gray and Leary (7) in 1935, by Lorge (8) in 1939, by Flesch (9) in 1943, and by Dale and Chall (10) in 1948. In each instance a

multiple regression formula was developed relating a criterion and some internal indications of expressional difficulty.

Perhaps the greatest departure from the formula of Vogel and Washburne is in the nature of the criterion. For them, it was the median paragraph-meaning score of those who had read and liked the book. Ojemann, on the other hand, used as the criterion a score indicative of the average reading ability of his subjects. Essentially, such a reading ability score involved the preparation of questions on each passage (used in the devolpment of the formula) and expressing it in terms of some well established reading test as a medium for reference. By this procedure each passage had a difficulty value or score which could serve as the criterion. This method, also utilized by Lorge, by Flesch, and by Dale and Chall, has as a criterion a measure of reading comprehension for each specific passage. At first glance, this seems to be the best possible criterion. Yet, its obvious excellence masks its fundamental weakness.

The criterion is the amount of comprehensibility in a given passage, measured by asking questions about the content of the passage. The questions are designed to reveal the reader's general understanding of the text, his grasp of specific details, his utilization of the ideas, and so on. The comprehension of a text, therefore, is measured, in part, by the response to the questions set for it. Such questions may vary not only in the level of language used, but also in the level of concepts considered.

If the questions are couched in a vocabulary less common than that used in the passage, fewer of them will be answered correctly; or if the questions require inferences that only well-informed or mature persons could make, few will be answered correctly. On the other hand, if the questions are expressed in simple language or if the questions are about easy and common ideas, more of them will be answered correctly. The number of correct answers to questions, therefore, must depend on the kind of questions set for the reader. The quality of comprehension a passage gets from its reader, moreover, is changed by requiring him to answer set questions. Whenever a reader is asked to read to get answers, he tends to read with greater care and precision.

The net result is that the procedure for measuring the comprehension of a passage influences the rating of it. The criterion score, at best, is an amorphous mixture of expressional and conceptual difficulties not only in the text but also in the questions framed to evaluate it. Lorge (8) estimated the so-called readability of 120 different passages by means of the Gray-Leary readability formula; he also estimated the readability of each group of questions designed to measure passage understanding. For each passage, therefore, Lorge had two scores: a Gray-Leary index for the difficulty of the passage and one for the difficulty of the group of questions about it. The relationship between the score for passage and the score for its questions was so low that it is reasonable to infer that only a third of the factors accounting for passage difficulty were present in the questions. The difficulty measure for a text, inevitably, is tied to the quality of the questions used to appraise it. The indeterminacy of the interaction of questions with

text, perforce, contaminates the criterion with ambiguity.

There is great confusion about the meaning of reading comprehension scores themselves. As far as is now known the difference between the reading comprehension in an average third, and in an average fourth, grade is much greater than the difference between the reading comprehension in an average tenth, and in an average eleventh, grade.

Reading performance in the lower grades represents more the sheer mechanics of the reading process whereas the scores in the upper grades represent more of the conceptual mastery of ideas. The units in which scores on a reading test are expressed fail to be either uniform in size or correspondent in process.

In *The Right Book for the Right Child* (11), in a discussion of books for the ninth grade and higher, the point is clearly made. "The lines of demarcation in the grading of books grow much less sharp as one reaches the higher grade levels. The books in this list may, therefore, be considered as somewhat more difficult than books in the preceding lists."

The measure of the readability or comprehensibility of a passage is not precise. The score is an undifferentiated mixture of different reading processes, unequal units, and indeterminacy. Nevertheless, the reading score allows broad, perhaps, too broad, differentiation among texts.

In most of the formulae the kinds of books or passages analyzed also put restrictions on the generality of application. As has been suggested. Vogel and Washburne base their formula primarily upon children's voluntary reading of fiction or fictionized material. Ojemann (5) used material ostensibly written for parents. Dale and Tyler (6), and later Dale and Chall (10) sampled material dealing with personal health. Flesch (9), using the identical material developed by Lorge (8), worked with reading exercises from the McCall and Crabbs *Standard Test Lessons in Reading*, a series of practice exercises in reading comprehension. The variety of the material, therefore, is constrained.

Most studies are based on children's reading. Except for Ojemann's research, no objective evidence about really *adult* material is available. For adult material, the criterion usually is the judged difficulty of books or of magazines. While such judgments are valuable, they tend to reflect factors not only of structural and conceptual difficulty, but also of the judge's interest in, and attitude toward, the content and its vehicle.

The various criteria of difficulty, impure though they are, are interrelated. Despite the fact that the prediction of any one measure, to a degree, gives evidence about other criteria of difficulty, the relationship is far from perfect. At best, there will be an approximate ordering of difficulty even though it will be unclear whether the order is a function of voluntary reading, tested comprehension, interest maturity, attitude, or some other considerations.

None of the formulae thus far developed gives adequate consideration to the concepts involved in the texts. The primary consideration has been given to *elements* of expression. A systematic review of readability formulae suggests that only four kinds of elements have been considered. These four are:

1. *Vocabulary load*, usually appraised in terms of vocabulary diversity or in terms of vocabulary rarity. Diversity is evaluated as the percentage of different words in a passage. Rarity is measured in many different ways: the relative number of uncommon or hard words, or the percentage of words not known to children at specified grade levels, or the relative syllables per hundred words. Diversity and rarity are, of course, related; and, obviously, the various measures of rarity are just different ways of measuring the same element.

   In general, the best single element for the prediction of any aspect of expressional difficulty is vocabulary load. The other elements add somewhat to the prediction, but generally not too much.
2. *Sentence structure*, usually appraised in terms of sentence length. The longer the sentence, the more probable that it is complex or perplexing. Each formula, whenever it uses an element in addition to vocabulary load, includes some measure of sentence structure, such as the average number of words per sentence, or the number of sentences that are simple, or some combination of such observations.
3. *Idea density*, usually appraised in terms of the number of different nouns or verbs, or occasionally of the number of abstract and concrete words, or of homely and recondite words. The relative number of prepositional phrases, sometimes used to evaluate idea density, also gives evidence about sentence structure.

   In general, the greater the relative use of prepositional phrases, the denser is the ideational content and the more complicated the style of the writer. In practice, however, the number of prepositional phrases is likely to be inexactly counted because many teachers and research workers do not know what a prepositional phrase is. In the sentence "John has to go to school" they, too often, count two prepositional phrases.
4. *Human interest*, usually evaluated in terms of the directness of approach. Among the devices utilized have been the relative number of personal pronouns, or words indicative of human interest, or words associated with early learning, or with homely experiences. As the human interest element increases, the passage's difficulty decreases. The other three measures, however generally indicate that difficulty increases with additional amounts of each.

No other internal elements of comprehensibility have been found useful in estimating passage difficulty. It is quite probable that, considering the ambiguity of the criterion, no other factor would add substantially to predicitions of expressional difficulty. Readability formulae have at least two fundamental weaknesses: they do not directly evaluate conceptual difficulty, nor do they consider the way in which the text is organized. The concepts within a passage may be involved and abstruse because the ideas are remote from experience, or because they are inadequately explained, or because they are extraordinarily abstract. Many history texts, for instance, suffer from the nature of the abstractions rather than from the level of vocabulary.

It must be emphasized, however, that of the four factors of expressional contributors to readability, vocabulary load is the most important. Some researchers, as a matter of fact, have found that an estimate of *vocabulary load*, in and of itself, is a sufficiently sensitive index of readability. G. A. Yoakam, for instance, uses a formula (unfortunately, as yet, unpublished) based only on a weighted value of the known frequency of occurrence of the words used in the passage.

Lorge (8) invesigated such a measure of vocabulary load. He gave a value of 1 to words that were in Thorndike's commonest one thousand words, a value of 2

to words that were in the second thousand most frequent words, and so on to the value of 20 for the words that were in frequency position 19,001 to 20,000. The correlation between that weighted index and the number of hard words (not in the Dale list of 769 easy words) (18) was so high that either measure of *vocabulary load* could have been selected for a *single-factor* estimate of readability. In the development of the multiple regression formula, however, *hard words* proved to be a somewhat better predictor than the *weighted index* when combined with the other factors of expressional difficulty, e. g., *sentence length* and *prepositional phrases*.

The weighted index of *vocabulary load*, whether that of Yoakam or of Lorge, is probably the best measure of passage difficulty for texts planned for children in grades below the fourth or fifth grade. Nevertheless difficulty based only on vocabulary frequency can be dangerous. Knowing that the word *run*, or *set*, or *mean* is a word of high frequency allows people to forget too easily that most high frequency words have a large number of different meanings. In the *American College Dictionary*, published by Random House in 1947, *run* has 104 numbered definitions, *set* has 67, and *mean* a mere 24. High frequency may, in general, indicate ease. Yet think of the difficulty of *run* in sentences like "I'll take a run in the car," "I've had a run of bad luck lately," or "When I get to the run, I'll ford it," as compared with its ease in "The dog runs after its master." Word frequency is not necessarily inadequate as a first order approximation; it fails to give consideration to other sources of textual difficulties.

For instance, all formulae neglect the organization within the passage itself. The second, and perhaps more important, source of text difficulty is the way the text is put together. Donald R. Murphy (14), the editor of *Wallace's Farmer and Iowa Homestead* has demonstrated that the arrangement of an article may be even more significant than some of the measured aspects involving merely expressional elements. He tried to find out about how many people read differing versions of an article. For example, here are two versions of a brief article on "What will you do with your corn?":

Version A

The method of handling the present corn crop may affect farm profits this year and have some influence on farm profits for a year or two ahead.

There are several different ways in which farmers might market their 1946 corn. The first way is to turn every possible bushel on the market this fall, keeping only enough corn to feed until new oats come along next summer.

Another method is to feed hogs to heavy weights, raise more early spring pigs, and feed your own cattle or buy cattle to feed.

Still another method is to put corn in storage on the farm selling just enough to pay farm expenses, or sell the corn and take a government loan.

Version B

What will you do with the big corn crop? The answer may tell you how much money you will make this year and even next year. Here are suggestions:

1. Turn every possible bushel on the

market this fall. Keep only enough corn to feed until new oats arrive next summer.

2. Feed as much as passible of the corn to hogs. Raise early spring pigs to eat more of this corn crop. Buy and feed cattle.

3. Store all the corn on the farm that the cribs will hold. Sell just enough corn to pay expenses. Or, sell corn for ready cash.

Version A does have longer sentences, more uncommon words, and fewer personal references than does version B. The more important difference, however, is that version A rambles in the corn whereas version B arranges it in a way that the reader can get at it and do something with it. Version B appeals to the reader's motivations and organizes the content clearly. It gets read.

Appeal in the sense of interest, and structure in the sense of organization or arrangement, perhaps, are even more important than elements of expression. Neither, however, has been adequately considered in the available readability formulae, Morover these formulae give little, if any, attention to the vividness of imagery or to the emphasis of dramatic expression; nor do they use the motivations within the reader, or consider his maturity level. Children in the first two grades probably would find animal stories, nursery rhymes, and fairy tales appealing, but children in the seventh and eighth grade would seek adventure, romance, and even humor in their reading. For adults, the range of appeals would be different and wider.

Despite the vagueness of the criterion and the restriction of elements involved in predicting it, the various formulae, if carefully used, will allow the classification of materials into broad relative order. The revised Washburne and Morphett formula (14) will permit the classification of children's voluntary reading. The revised Lorge formula (16) will give some evidence on the comprehension difficulty of the kind of reading pupils have to do in school, Ojemann's formula (5) will permit the evaluation of parent education material, the revised Flesch formula (17) will estimate comprehension difficulty of magazine material, and the new Dale and Chall formula (10) will make for adequate appraisal of health materials. There will be some correlation between the rank orders given the same material by different formulae.

When the formulae are used for estimating a single aspect of difficulty, they are used correctly. A device such as a readability formula, regrettably, almost begs to be misused. Since it promises the layman a measure of readability, too many are misled into applying the formula as a panacea for inadequate writing. One person, advocating the formula as a rule for writing recommended that the sentence "I am going to town" should be rewritten as "I am going townwards." She explained that this would reduce sentence length, involve fewer different words, and eliminate one prepositional phrase. Thus, she explained, the sentence would become easier to read. The dependence upon the formula as a guide for writing may result in a stilted and primerized style; it may throw away the ideas with the formulaic bath.

Every writer who knows his craft recognizes that his message should be couched in words that will be understood, and in

sentences that reveal rather than hide its meaning. He also knows that if he tries to compress his ideas into few words he may succeed in communicating nothing but confusion.

The style of writing helps make the book interesting, popular and understood. James Earl McClintock (15), in his master's degree essay, studied three versions of a single publication. He found that "The vocabulary, the sentences used, and the handling of the paragraphs constitute the style or manner of expressing the ideas of the writer. This style, to be effective, must be adapted to the reader and the subject [topic] . . .Words of concrete meaning create positive impressions and hold the interest longer than do words of a general or an abstract nature. . .Short sentences make for rapid reading and add emphasis to the writing. Long sentences aid in grouping details and in summaries. Many long sentences slow down the reading and a repetition of many long or of many short sentences make monotonous reading. A mixture of short, medium and long sentences make easy reading and popular bulletins." Variety adds zest to writing as it does to life.

Readability formulae are no panacea. They do not tell anything about the kind of ideas expressed or the interrelationships among them. At best, they are yardsticks. If they are not inflated into a recipe for writing, they are a useful adjunct in the objective evaluation of written and spoken materials. Their use, however, cannot be a complete substitute for the wisdom of experience.

BIBLIOGRAPHY

1. Vogel, Mable and Washburne, Carleton. "An objective method of determining grade placement of children's reading material." *The Elementary School Journal* 28: 373-381: January 1928.
2. Washburne, Carleton and Vogel, Mary. *Winnetka Graded Book List.* American Library Association, Chicago, 1926.
3. Thorndike, Edward L. *The Teacher's Word Book.* Teachers College, Columbia University, New York, 1921.
4. Lewerenz, Alfred S. "Measurement of the difficulty of reading materials." *Educational Research Bulletin Los Angeles Public Schools* 8:11-16: March 1929.
5. Ojemann, Ralph H. "The reading difficulty of parent education materials." *Proceedings of the Iowa Academy of Sciences* 40: 159-170: April 1933; see also "The Reading Ability of parents and factors associated with reading difficulty of parent education materials." Researches in Parent Education II *University of Iowa Studies, Studies in Child Welfare* 8:11-32: March 1934.
6. Dale, Edgar and Tyler, Ralph W. "A study of the factors influencing the difficulty of reading materials for adults of limited reading ability." *The Library Quarterly* 4: 384-412: July 1934.
7. Gray, William S. and Leary, Bernice E. *What Makes A Book Readable.* University of Chicago Press, Chicago, 1935.
8. Lorge, Irving. "Predicting reading difficulty of selections for children." *Elementary English Review* 16:229-233: October 1939.
9. Flesch, Rudolf. *Marks of Readable Style* Bureau of Publications, Teachers College, Columbia University, New York, 1943.
10. Dale, Edgar and Chall, Jeanne S. "A Formula for Predicting Readability." *Educational Research Bulletin* 27: 11-20: January 21, 1948: and "A Formula for Predicting Readability: Instructions" *Educational Research Bulletin* 27: 37-54: February 17, 1948.
11. *The Right Book for the Right Child.* New York, The John Day Company, 1942. Third edition, completely revised.
12. McCall, William A. and Crabbs, Lilah M. *Standard Test Lessons in Reading.* Books II, III, IV, V. Bureau of Publications, Teachers College, Columbia University, New York, 1926.
13. Murphy, Donald R. "How plain talk increases readership 45% to 66%." *Printer's Ink*: 220: 35-37: September 19, 1947.

(Continued on Page 25)

The Use of Vocabulary Lists in Predicting Readability and in Developing Reading Materials

E. W. DOLCH[1]

Teachers everywhere are much interested in determining the "readability" or reading difficulty of materials presented to them by publishers. After all, everyone knows that children should not be asked to attack material that is too difficult for them, for fear they may lose their interest in reading or develop bad reading habits. Research departments of many school systems are interested in reviewing new reading materials and determining their reading difficulty so that the teachers in the system may be advised of it. All of these persons look to research studies to tell them how to determine readability for different grade levels.

Each of the various studies on readability uses a vocabulary list as a fundamental part of its method. A sampling of the new reading matter is compared with this vocabulary list. A conclusion is reached as to the difficulty of the reading matter. The process itself is chiefly a clerical one, but the interpretation of the result requires a considerable understanding of the reading process and of the characteristics of word lists. In short, there are a number of precautions one should take in making such an interpretation.

To aid teachers and others who wish to make use of vocabulary lists in determining difficulty of reading matter, we shall discuss a number of the problems that should be kept in mind. We shall refer to various vocabulary studies as they are related to the problems dealt with. A list of these studies appears at the end of our discussion.

1. *Vocabulary difficulty is a basic element in reading difficulty.* We may say that the understanding of words in reading is basic because without understanding of word meanings there can be no reading. Words are to reading matter what bricks are to a house. When you look at the house, you do not think of the bricks; but without the bricks there would be no house. We do not read words; in fact, we may not be conscious of single words at all as we read. But the reading matter is after all made up of words, and without word meanings there would be no sentence meanings and no paragraph meaning.

2. *Vocabulary difficulty is only one part of reading difficulty.* It is true that if a child does not know the meanings of the words he sees he cannot get the meaning of the whole. But we cannot turn that statement around and say that if he *does* know the meanings of all the words, he can fully understand what he reads.

In the first place, the reading of every sentence is a test of the span of attention. Every sentence includes the meaning of each word and also the relationship of each

[1]Professor of education, The University of Illinois.

word with the other words. The reader begins at the capital letter and glances along the line. He "takes in" more and more word meanings, more and more word relationships. How much can he "take in" and still keep the thought straight?

It is true that some sentences are so built that one can read them "in installments," so to speak, understanding a part before going on to the next part. Other sentences cannot be understood until the very end is reached. But in either case the reading of the sentence is a test of the reader's span of attention, of his ability to put ideas together and to get a single unified thought as a result. The longer the sentence, the more difficult it is to get this unified thought. So mere sentence length is a factor in reading difficulty, over and above word difficulty.

Unusual word order or order of sentence parts is also a factor in reading difficulty. Some proverbs are a case in point, as in the one, "All is not gold that glitters." The reader may know the meaning of every word but not have the faintest idea of what the sentence says. The clause "that glitters" is put out of place by the author for effect. The reader would think it modifies "gold" and say, "Of course gold glitters, and of course all is not gold. But why say such a thing?" Not all sentences have such unusual order of parts, but any unusual order introduces a real factor of reading difficulty.

Another factor, which might be called "idea difficulty," is of greater importance than either of the above. "Idea difficulty" means the degree of remoteness of the idea from the reader's past experience. The word meanings may be familiar but they may be put together to make a statement that has little relation to the reader's experience or thinking. This factor has been little studied, except perhaps in a count of abstract terms, but it is of first importance in reading matter that concerns fields of experience and thought into which children have had little chance to enter.

3. *Consider the problem of multiple meanings.* In all determining of grade placement by word difficulty, an error creeps in because of multiple meanings. This is the more true as we go up in the grades. We come to the "square of a number" and the vocabulary list says the words are known, since a child knows what the shape of a square is in drawing. We come to the "root of the trouble" and "root" is said to be known because children know the root of a plant. In fact, almost all the common words sooner or later get derived and figurative meanings that present word lists do not properly make allowance for.

One book publisher does have a "graded meaning" list, and Dr. Lorge has meanings but which is not available to the general public because of cost. Here is, therefore, a situation in which the user of word lists must fall back on his own ingenuity. One device is to say that "every word on the list is assumed to have its commonest meaning. All other meanings are not on the list." Such a plan requires judgment by the user of the list and it will not be entirely fair either. In the first place, the maker of the list intended to include more than just one meaning in many cases. Second, children can often derive uncommon meanings from common ones as they read. So taking multiple meanings

into account is a difficult problem. Probably in the future the publishers of lists will append after each word form the meanings intended to be included. Until that time, most persons will use the words as given, ignoring the problem of multiple meanings.

4. *Meaning vocabulary and sight vocabulary are very different things.* All vocabulary lists are assumed to be "meaning" lists. That is, they assume that they are measuring whether the child will know the *meanings* in the new material that is being studied. If the meaning is on the list and the meaning is in the book, it is assumed that there is no difficulty. However, this assumption also assumes that the child, when reading, can *recognize* the word, either by sight or by sounding. The list for instance, may include "ability" as an easy word, the book may use "ability," and the child may know the meaning of "ability," but if the child does not know that the word *says* "ability" and cannot find out what the word says, it is an unknown word to him.

We find this same trouble with vocabulary tests. A child may actually have an eighth grade meaning vocabulary but still score fourth grade on a standard test just because he cannot recognize the words. They are "unknown" words to him because he cannot tell what words they are. As a result a "meaning vocabulary" test becomes instead a sight or sounding vocabulary test.

Books for the first three or four grades are likely to be written within the area of familiar words which may be known at sight by children. So at that level, meaning vocabulary and sight vocabulary may agree. When we get above these grades, we find many words that are not likely to be in the sight vocabulary. These words must then be sounded to be recognized. If we compare the reading matter with a list, it may be said to have a seventh grade meaning vocabulary and therefore supposed to be suited to seventh grade or higher. But if the children have not been taught sounding, this same book will be far over their heads because they cannot use their meaning vocabulary in reading it.

5. *Choose the size of list to fit your needs.* In using a word list to check the vocabulary of new material, we are finding out whether words are "on the list" or "not on the list." Here the size of the list makes a very great difference. There is always the temptation to use as small a list as possible for ease in checking but that is not the main consideration.

A little thought will show that for easy material intended for the lower and middle grades, a small list may do well enough. Most of the words in easy material will be on the small list. So the difference in difficulty comes in the words "not on the list," that is, in the smaller number of "hard words." With harder reading material, however, the small list does not work so well. In checking the harder material, we find a greater and greater number of words not on the list, and there is no way of telling how many of those words are really hard. They are "beyond the list" but how far beyond?

For the upper grades and high school therefore, and for adult material, a longer list is needed. For such material, only a long list will tell the differences between easy and hard words. Conversely, if too long a list is used on easy books nearly all

the words will be on the list, and so one book will seem not different from another.

An arbitrary suggestion might be that up to Grade VI a list of 1000 or 1500 words will do, but after that a list of 3000 or more is needed. A study needs to be made of just this point, using lists of various sizes on various levels of material to determine the best relationship between size of list and difficulty of material.

If the Thorndike list of 10,000 words is used or the Thorndike-Lorge list of 30,000, the words in the new material would fall into any of the different thousand lists or beyond them. If you use the Rinsland list with different grade levels, the words will fall into different levels or beyond them. Then the problem is to evaluate different percentages at different levels. No one has worked out a plan for doing this satisfactorily.

6. *Study the words which are "not on the list."* If the user of a list will write out and examine the words which are not on the list, he will see several things which are significant in the checking for reading difficulty.

First of all, he will probably agree that most of the "not on the list" words are obviously harder than those on the list. This fact will strengthen his faith in the particular list and give him greater confidence in using the results. He will be using subjective judgment, but the user of any vocabulary list is going to have to make some subjective estimate of its value. He should get that estimate by study both of the list and of the words it excludes.

On the other hand, study of the words may give the decided impression that they are in many cases no harder than many of the words included in the list. Is this a defect of the list? Usually not. The reason is that the list endeavored to cover a certain level of difficulty but did not fully succeed in doing so. A list that includes the "thousand commonest" words does not get absolutely all the words of that degree of "commonness." Many are missed in the process. It is the same with a list that is intended to show the 3000 commonest, or the 10,000 commonest. There are no methods of tabulation which will completely cover any certain level of difficulty. Therefore in the case of the particular reading material being studied, the author may have skillfully sensed the level of difficulty required and kept to that level. He, therefore, will have used many words that were appropriate but that had been missed by the list.

For these reasons, it is wise to see if all the words "not on the list" are actually "beyond the list" or are more nearly parallel with it in difficulty.

7. *Consider the source of a list.* When you wish to determine the readability of a book you should ask, "readability for whom?" Then you need to consider if the list you wish to use fits your purpose.

If children are to read the book you are considering, you need a list that gives you words familiar to children. If you are thinking of the average adult, you need to consider whether the list gives you words familiar to the average adult. Therefore scrutinize carefully the source of any list used. Sources of the lists commonly used are given in the description of the lists which is at the end of this discussion.

This caution, to consider the source,

was first suggested by the first widespread use of the Thorndike list of 10,000 words. Dr. Thorndike said very clearly that his list showed the most frequent words found in adult reading matter. He listed the reading matter, and it showed a heavy weighting of classical reading materials. Then people made the general assumption that the most common words known to educated adults would be the words known to children. They ignored the fact since called to our attention, that adults and children live in rather different worlds.

It is true that the "service words" of the language—the simplest verbs, adverbs, adjectives, conjunctions, and the like—are used both by children and by adults. But beyond that common element of daily speech, the experiences are in many ways very different. For instance, the word "puppy" and other names for small animals are very common to children but are relatively uncommon in use in the adult world. "Puppy" is listed by Thorndike as in the 5th thousand.

Similarly the mistake has been made of assuming that the limited vocabulary of Basic English, devised by Professors Ogden and Richards, would make a good vocabulary for children. A comparison of Basic English with children's lists shows that about half of Basic English is not at all common in children's usage. This is natural, since the authors of Basic English purposely selected abstract adult words, such as "instrument," as substitutes for large groups of concrete words such as "shovel" and "rake."

Some lists are based on words which children *use*, and one must ask if *usage* corresponds in any very close way with *children's knowledge*. So much of what children know they have no reason for talking about. They know what a carpenter is but they have no reason for talking about carpenters and carpenter work. Children know much about adult life, but largely as spectators of it. If reading is to fit children's *knowledge*, it must fit something more than just children's usage of words. A list to be used on children's reading matter should present child knowledge of words as well as child usage of words.

No doubt at some time we shall have better vocabulary lists, ones which do more than just roughly correspond to different large experience levels. We may at some time have lists for each grade showing the "characteristic grade experiences" of children. Meantime we must use as intelligently as possible the lists we have. Their source and method of derivation tell us much about them.

8. *Special subject matter lists must be considered in some cases.* It is a well known fact that different fields of interest have different vocabularies. For instance, it has been found that different units studied in school imply different special vocabularies. It is also recognized that special subjects such as health, science, arithmetic, and so on have special vocabularies.

Up to now, special vocabulary lists have not been used to determine "readability" because there are no norms for technical or special subject matter books. The lists have been used to discover just how heavy a load of technical words a certain book may have, but then no one knows whether that load is enough or too much. For instance, no one knows just what the average load of arithmetic vocabulary

should be at any grade. Therefore, in the special subjects, the special vocabulary is usually considered a matter of curriculum planning not of readability. If there seem to be too many technical terms, we do not say the book is unreadable but that it is hard to teach.

Obviously we need a study of "readable" books in the special fields. The only difficulty is that teaching a subject, we make books on that subject readable at the level at which we teach it. Therefore "readable" in a special field must mean "readable after a certain amount of teaching."

Developing Readable Materials

In the preparing of materials for certain levels of reading ability, the use of word lists is very important both because word meanings are so basic to understanding in reading and because it is so hard to measure other types of reading difficulty. Most writers and editors turn readily to word lists. It is worthwhile, therefore, to discuss for a moment methods that are or may be used in using word lists in the preparation of reading materials, together with certain cautions with regard to them.

One plan that seems to be followed by those preparing readable materials is just to take any reading material they find and wish to use, check the vocabulary with a list, and "substitute easy words for hard ones," thus securing the desired grade rating. This is a good plan as far as it goes. It does not adapt method of presentation. It does not adapt language pattern. It does not consider that merely substituting words for words may give inaccurate meanings or may damage the idiom of the language.

A second plan seems to be to have someone "rewrite" the material in the easier vocabulary. Such a "writer" tries to tell the story or explain the idea in easier words, with one eye always on the list to make sure he keeps within it. The result we see in many school readers and other books: It is often a very "lame" or "wooden" kind of writing, shorn of all spontaneity and creativeness. It is this kind of writing that has made so many teachers hate the word "adapted."

A third plan, and one that is the most successful, starts with the writer studying the audience he is writing for. If it is the average adult, he sits in buses or restaurants or wherever people talk, and listens to their kind of language and their type of vocabulary. If the writer is writing for children of a certain grade, he sits in the grade room of various schools for several days and follows the children to the playground, listening to how they express themselves. Then the writer, when he has the "feel" of his audience, sits down to write, keeping the audience right before him in his mind's eye. He writes directly to them. After he has written, he may actually read the material to the desired audience and watch their reaction or talk with them about it. Then, when he thinks he has "hit it right" he takes out his word list and figures his percentage. If it is wrong, he may alter a word here or there, but he will not need to do much. He will have found the level of his vocabulary and language pattern beforehand. He will have *created* at that level, not just "adapted" to a level.

Perhaps all three of these methods of preparing material will have to be used. For adults, perhaps the more mechanical

adapting may do for many purposes. For children, it is to be hoped that the creative method will be used. Reading matter for children needs the charm and appeal that can only be secured by creating at their level.

Lists of Vocabulary Studies.

1. Cole, Luella, *The Teacher's Handbook of Technical Vocabulary*, Bloomington, Illinois: Public School Publishing Co. 1940.

"The most widely used textbooks in each subject were first gone over by readers who were instructed to list all words occurring in the texts read which, because peculiar to the subject or uncommon, might be difficult for the children in the grades in which the subject is taught. . . .The list for each subject was then sent to teachers of that subject. . .to check those words they considered absolutely essential . . .important but not essential. . .or unnecessary. . . .Only four subjects were rated by less than 35 teachers."

2. Dale, Edgar, "A Comparison of Two Word Lists," *Educational Research Bulletin*, Ohio State University, Columbus, Ohio; Vol. 10 pp. 484-489, Dec. 1931.

"The Dale List of 769 Words is made up of words which are common to Thorndike's First Thousand Most Frequent English Words and the first thousand most frequent words known by children entering first grade."

3. Dale, Edgar and Jeanne S. Chall, "A Formula for Predicting Readability," *Educational Research Bulletin*, Ohio State University, Columbus, Ohio, Vol. 27, No. 1, Jan. 1948.

"Our word count was based on the Dale List of approximately three thousand words. This list was constructed several years ago by testing fourth graders on their knowledge in reading of a list of approximately ten thousand words. This larger list included the most common words in the Thorndike, Buckingham and Dolch, and other word lists. Words such as *milkman, carrot, candlestick, catbird*, and so on which appeared in the high thousands on the Thorndike list were also tested with fourth graders to see whether they knew them. An attempt was made to include all words that fourth graders would possibly know. A word was considered known when at least 80% of the fourth graders checked it as known."

4. Dolch, E. W., "Graded Reading Difficulty," being Chapter XXI in *Problems in Reading*, Champaign, Ill: Garrard Press, 1948, together with Chapter X of the same book "The First Thousand Words for Children's Reading" (Work sheet containing the thousand words also available)

"The Dale list of 755 words was increased to 1,000 by additions from the Interview Vocabulary study, a list of words known to 75 children out of 100 entering first grade, the additions being made on a basis of a topical analysis of the Dale list and the filling of gaps in the various topics."

5. Eaton, Helen S., *Semantic Frequency List for English, French, German, and Spanish*: A correlation of the first six thousand words in four single-language frequency lists, Chicago, University of Chicago Press, 1940.

6. Fries, Charles C. and Traver, A. A. *English Word Lists*. Washington, D. C.: American Council on Education, 1940.

The authors give a complete history of word counts and limited vocabularies with notes on method of choice, dictionary counts, inadequacy of mere frequency of use as a determiner of significance, and the need for semantic counts.

7. Gates, Arthur I., *A Reading Vocabulary for Primary Grades*. New York: Bureau of Publications, Teachers College, Columbia University, 1935.

"The words were originally selected from

the following sources: 1. the 2,500 words of highest frequency as determined by Thorndike's count... 2. any words not in the 2,500 from Thorndike found in the thousand words of highest frequency as determined by a count of words in a selection of children's literature. 3. all additional words found in the thousand most frequent words in a series of readers for the primary grades... 4. all additional words found in the thousand most frequent words in the spoken vocabulary of young children. Each of the 4,300 words was appraised for merit for use in reading at different stages during the primary grades on the basis of utility, interest and difficulty by judgment of experts." (list revised in 1935)

8. Lorge, Irving, "The English Semantic Count," Teachers College Record Vol, 39, pp. 65-77.

"The materials used for the count were selected as a sample of journalistic, learned and recondite, adult fiction, textbook, juvenile fiction, quotations, and adult non-fiction reading. The material to be counted has approximately five million running words."

9. Lorge, Irving, and Thorndike, Edward L., *A Semantic Count of English Words*. New York. Institute of Educational Research, Teachers College, Columbia University, 1938.

A semantic supplement to the Thorndike word list. (See 12, 13)

10. Rinsland, Henry D., *A Basic Vocabulary of Elementary School Children*. New York: Macmillan Co., 1945.

"The letters addressed to school officials (of 1500 selected schools in all kinds of geographic, economic and social areas) stated that all kinds of children's writings, representing their freest and most natural compositions, were desired. ...personal notes, stories, poems, compositions in many school subjects, examination papers in nontechnical subjects, articles for school papers that were not corrected by teachers, and reports on projects, trips, and observations....100,212 compositions in the eight grades were used. Only one composition from each child was used. ...The words of Fry and Trent....furnished 4,630 pages of conversation material (from Grade I)"

11. Stone, Clarence R., *Stone's Graded Vocabulary for Primary Reading*. St. Louis, Mo.: Webster Publishing Co., 1941.

"In compiling and grading the vocabulary of 2,000 words contained herein, various previous vocabulary studies of primary reading books, the Gates revised list of 1935, and standard lists of words most commonly appearing in children's spoken vocabulary have been taken into account. In addition an extensive vocabulary study of twenty-nine primers, twenty-seven first readers, twenty second readers, and eleven third readers was made. The words are graded mainly on the basis of the placement of the word as a new word, as a rule, in the various series of readers."

12. Thorndike, E. L., *A Teachers Wordbook of 20,000 Words*. New York: Bureau of Publications, Teachers College, Columbia Univ. 1931.

"About ten years ago I published a list of 10,000 words which were found to occur most widely in a count of about 625,000 words from literature for children, about 3,000,000 words from the Bible and English classics, about 300,000 words from elementary school textbooks, about 50,000 words from books about cooking, sewing, farming, the trades and the like, about 90,000 words from the daily papers, and about 500,000 words from correspondence. Since then I have made counts from over 200 other sources .....As a result I am now able to extend the list to 20,000 words and to revise and improve the selection of the most important 10,000..."

13. Thorndike, Edward L., and Lorge, Irving, *The Teacher's Word Book of 30,000 Words*. New York: Bureau of Publications, Teachers College, Columbia University, 1944.

"This book is a greatly improved extension of the Thorndike *Teacher's Word Book*, published in 1921, and of the extension of it to include 20,000 words published in 1931. It includes the data of these two counts and also of three other counts of over 4½ million words each. It enables a teacher to know not only the general importance of each word so far as frequency of occurrence measures that, but also its importance in current popular reading for adults, as shown by the Lorge magazine count, and its importance in such juvenile reading as schools and libraries approve."

READABILITY FORMULAE - AN EVALUATION

(Continued from Page 16)

14. Washburne, Carleton and Morphett, Mabel Vogel. "Grade placement of children's books." *The Elementary English Review* 38: 355-364: January 1938.
15. McClintock, James Earl. *What Makes a Popular Bulletin Popular*. Master's essay, University of Wisconsin, Madison 1929.
16. Lorge, Irving. "The Lorge and Flesch Readability Formulae: A Correction." *School and Society*, 67: 141-142: February 21, 1948.
17. Flesch, Rudolf. "A New Readability Yardstick." *Journal of Applied Psychology*, 32: 221-233: June, 1948.
18. Dale, Edgar. "A Comparison of Two Word Lists." *Educational Research Bulletin* 10: 484-498; December 9, 1931.

Typography and Readability

HAROLD E. BURTT[1]

Introduction

The preceding chapters have discussed readability primarily as influenced by content—the vocabulary or ideas involved. Another factor which contributes to readability is more mechanical,—the typography. This is not primarily an educational problem, but the educator sometimes does have latitude in furnishing specifications to the printer. Hence it is well to round out the discussion of readability by noting some of the common variables in typography.

Methodology

Photographs of eye movements. In the approach to any scientific problem the methods used are just as important as the results obtained, because results based on inadequate or faulty methods are meaningless. Rather than describing methodology in connection with each experiment cited below, it will be well, at this point, to make brief mention of the principal techniques employed in evaluating legibility Eye movements play an important role in reading and while it is possible to sit opposite a reader and look across the top of the book to note how his eyes move, this method is inadequate for scientific purposes. A better technique projects a beam of light upon the moist surface of the eyeball which reflects it onto a moving picture film. As the eyeball moves while reading, the reflected image moves back and forth on the film. It is thus possible to determine where the eye fixates and for how long and how rapidly it moves. Another technique attaches small electrodes to the skin in the region of the eyeball and records electronically the changes in voltage produced as the eye muscles function.

Speed of reading. Many investigations have used speed of reading as a criterion for the readability of printed material. Earlier efforts merely had the person read material at his "natural rate" and timed him with a stopwatch. This was none too satisfactory because it was impossible to determine how thoroughly he read the material. A better approach secures some indication as to whether he knows what he has read—what is conventionally termed "comprehension." For instance a standard series of short paragraphs has one wrong word in each and that wrong word is located near the end of the paragraph. The reader has to go through the set as rapidly as possible and mark the wrong words. A typical paragraph follows:

> "When Bill came downstairs to breakfast in the morning he found a long letter on the kitchen table. Before he did anything else he picked it up and ate it."

One must read the entire paragraph in order to locate the wrong word "ate."

It may be noted in passing that speed of reading reflects not merely typography but vocabulary level and general difficulty of the material. Fiction vs. a philosophical treatise would be a case in point. However, it should be pointed out that in experiments with this technique the matter of

[1]Dept. of Psychology, Ohio State University.

comprehension is essentially kept constant. Previous research has determined the difficulty of each paragraph and then two "forms" of the test are made up which are of equal difficulty. One form then can be printed in all capitals and the other in small letters and any difference in speed of reading found will reflect primarily the difference in typography.

Short exposure. The fact that in actual reading a person normally looks at a given word or group of words for only a fraction of a second suggests a technique of presenting material for extremely brief periods. One can look through an eyepiece controlled by a camera shutter set at, say, a tenth of a second. A more refined method controls visual adaption by providing a pre-exposure field illuminated to the same degree as the printed material. If exposure is held constant it is possible to determine the percent of the words that can be read in a certain typography.

Maximum distance. Another procedure is to present a letter or a word at such a distance that it is unrecognizable and move it toward the reader until he can just make it out. A specimen that can be read at a distance of 15 feet is more legible than one which has to be brought in to 10 feet.

Focal variator. A device which has been used in only a few laboratories has a series of lenses moving reciprocally so that by turning a crank the printed material may be thrown in and out of focus on a ground glass screen without changing the size or the brightness.

Blinking rate. There are some indications that when reading is difficult a person is inclined to blink more frequently. It is fairly simple to watch a reader and count the blinks per unit time, perhaps using an ordinary message counter. Photographic techniques are also possible. There is, however, some controversy among the experts as to the validity of this technique.

Significance of differences in scientific experiments. In many of the investigations to be described below the results indicate a larger proportion of people showing one trend rather than another or a larger average time required for one kind of performance than another. In such situations we must be very careful in drawing conclusions regarding the existence of real differences. The appropriate statistical formulae are beyond the scope of the present chapter but it will be in order to present the basic principles. Suppose we are investigating the comparative legibility of the same words when printed in small letters or in capitals using the speed of reading the short standard paragraphs. If we do the experiment with one person and find that he covers more paragraphs when they are printed in small letters, we would not be warranted in assuming that in general small letters are more legible because the next person that came along might produce just the opposite result. If we take fifty people and most of them favor the small letters, we are getting closer to the truth; if we use a hundred we are still more certain of the results. There presumably is a real relationship which we approximate by running our experiment with a limited group of people. It is a question of how typical is the sample of people which we examine. Obviously the more people we have in our sample the more closely we approach reality. Then in the second place if the people in our sample

agree with each other pretty well, the results take on greater significance. If almost everybody shows the trend in favor of small letters and only an extremely small minority shows the reverse trend, it is more probable that we have the real thing. The statisticians have not-too-involved methods for determining the extent to which a sample actually is typical of what would be obtained with a much larger sample or for determining what is called the "significance of the difference." What they finally come out with is an index which can be interpreted something as follows: "The chances are 99 out of 100 that if we did the experiment over again with another sample, we would still find the trend in favor of the small letters." When probabilities like that just stated, are sufficiently large, it is common practice to say that the difference is "significant." What this amounts to is that when reading scientific articles one must, perforce, be alert to considerations like the foregoing. A technical article which does not give some indication as to whether differences are significant should be regarded with some skepticism. In the following discussion when conclusions are drawn on the basis of differences, some indication of their significance will be included.

Type Face (Style)

The type founders have devised a great variety of type faces or styles or families. These are designated by trade names such as Cheltenham, Bodoni, Caslon, and a few examples follow:

14 point Cheltenham bold

14 point Bodoni bold

14 point Caslon bold

Close examination will indicate minor differences in serifs or in heaviness of the strokes, or in uniformity of width of strokes. Bondoni, for example, has some lines that are comparatively very light.

The results of investigations of the effect of such variables on legibility appear to depend on the methodological approach. On the one hand experiments with the maximum distance technique brought out the general principle that the most legible letters are those which have fairly heavy strokes and uniformly heavy strokes. For instance, take the capital letter N in a type face which has a heavy diagonal stroke and two light vertical strokes. When this is seen at a distance all the reader gets is the diagonal. This might equally well be part of V, M, W, or A and without the additional light strokes which at that moment are imperceptible one cannot judge correctly. An investigation of legibility of highway signs indicated that the best width of the stroke was about 18 per cent of the height of the letter.

On the other hand investigations using the speed of reading technique (short paragraphs) tended to minimize differences like the foregoing. Comparing one style (Scotch Roman) as a standard with half a dozen others the differences were of the order of two or three per cent and of little significance statistically. In everyday reading one does not look at every letter in the word but gets the general shape of the word and certain cues and landmarks and infers the rest. Evidently the variations in type face do not greatly influence this process. However, both methods show unfavorable results with ornate typography like Old English or Cloister Black. Here

the letters are so unusual and confused by curlicues that the reader cannot depend on the usual shape of the word or familiar landmarks.

Thus the influence of type face on readability depends largely on the conditions under which the reading takes place. From the standpoint of reading speed there is no very great difference between many of the type faces that are in common use and it is only when concern is with material near the threshold of visibility that this variable becomes very important. The schoolroom involves the former condition. It might be mentioned, however, that in an arrangement which combines several type faces in a single page difficulty does sometimes arise. This is more apt to occur in advertisements than in educational material. In one instance speed of reading tests showed that mixed type faces (with also some mixture of sizes) produced an 11 per cent loss in efficiency that was clearly significant.

Capital vs. Small Letters

There are occasions when material is printed all in capitals, usually for emphasis. When experimenting on this problem, it is advisable to use capitals and small letters all of the same "point," for example both of them in the size they would appear on an ordinary printed page.* Here again the results vary somewhat with the experimental technique and it makes a difference whether we are dealing with isolated letters or with words. With isolated letters and the maximum distance method the results are consistently in favor of the capitals for the obvious reason that many of them actually are larger than the corresponding small letters. When using words however instead of single letters and a short exposure method a typical group of persons read 43 per cent of the words that were printed in capitals and 56 per cent of those that were in small letters. Again, the speed of reading technique yielded a result in favor of the small letters of approximately 12 per cent. All the foregoing differences were large enough to be statistically significant and thus represent real trends. When eye movements were photographed during reading, all-cap material required on the average 12 per cent more fixations and the readers took in about 12 per cent fewer words per fixation. All this suggests then the superiority under ordinary reading conditions of material printed in small letters. The possibly increased emphasis or tendency to attract attention that might go with capital letters in certain situations such as advertising apparently is offset in ordinary reading by the inability to use the cues and landmarks by which we ordinarily read material without having to devote our attention to every single letter.

*Typographers figure approximately 72 points to an inch, that is 24 point type would set 3 lines to an inch and 12 point type would set 6 lines to an inch.

Length of Line

Experiments by the speed of reading technique indicate an influence of the length of the line. It will be most convenient to discuss length of line in terms of millimeters in order to avoid confusing fractions of an inch. There are approximately 25 millimeters to an inch. With ten point type one form of the reading test was in lines 80 millimeters in length, while line length in the other form ranged from 59 to 152 millimeters. At the ex-

tremes just mentioned the loss in reading speed was 6 or 7 per cent and statstically significant. With line length closer to the standard the differences were not as great although the 80 millimeter was consistently the best. Similar results were found when photographing the eye movements. Thirty-eight millimeters in comparison with 80 millimeters required 16 per cent more fixations; the reader took in 14 per cent fewer words per fixation, spent 8 per cent more time in each pause and the "perception time" was increased 25 per cent. Similarly, 180 millimeters in comparison with 80 required 8 per cent more pauses; secured 11 per cent fewer words per pause; the pauses were 4 per cent longer and the perception time was increased 14 per cent. Thus with 10 point type the optimal length of the line is somewhere around 80 millimeters. The experimenters in this particular case suggest between 75 and 90 as a range within which conditions are reasonably favorable.

Further question arises as to the optimal size of type with reference to the length of line. The foregoing experiment was confined to 10 point. In further experiments with the speed of reading technique the 80 millimeter 10 point was taken as the standard and the other form of the test was set line-for-line in different sizes of type, that is with the same words in each line. Thus, larger type would yield a longer line. The test ranged from 68 millimeters 6 point up to 115 millimeter 14 point. Here again the 80 millimeter 10 point proved to be most effective. The outside figures just mentioned were respectively 7 per cent and 4 per cent slower than the standard and the differences were all significant. To check one other aspect a further experiment kept the length of the line constant at 80 millimeters but changed the size of the type so that it was 8, 10, 12 or 14. The 10 point was superior to each of the others. Differences were 5 or 6 per cent and significant.

What this all adds up to is that a line of moderate length somewhere in the vicinity of 80 millimeters is the most favorable from the standpoint of legibility. Several explanations have been suggested to account for the difficulty with the long line. One is the tendency for the reader to get lost when his eyes come back to the beginning of the next line. Photographs of the eye movements actually show some fumbling at the beginning of long lines. Furthermore, when reading one line a person gets some of the content of the following line in the margin of his attention and these premonitions of meaning are helpful when he comes to the following line. If the lines are too long the premonitions obtained in this way are too remote from the present context.[1]

Space Between Lines

If type is set solid, that is with the bases on which the type is cast placed directly in contact with each other, the printed lines may be a bit too close together for effective reading. The descenders in one line interfere with the ascenders in the following line. Consequently it is common practice to set a little lead between the lines. This lead is gauged in points where 1/72 of an inch is one point.

[1]For an extensive reference summarizing much of the experimental material noted thus far and also other experiments, see Paterson, D. G., and Tinker, M. A. *How to Make Type Readable.* New York: Harpers, 1940, 209 pp.

The experimental technique most frequently used in this connection is speed of reading. It is necessary to relate amount of lead to other factors such as length of line or size of type and the procedure becomes complicated. It will be helpful, however, to cite enough data to indicate that in conditions which are apt to be encountered in ordinary books and magazines the addition of a little lead does help. For instance, with 10 point type and lines 80 millimeters in length the addition of one point lead made an insignificant change in the speed of reading, but two points caused about a 7 per cent improvement which was clearly significant from a statistical standpoint. Four point was only 5 per cent superior to the solid type. Apparently, it was not necessary to go that far and the use of unnecessary lead wastes paper and increases cost. By way of contrast, with 12 point type and a somewhat longer line the leading was not particularly advantageous. In the other direction, with 8 point type one or two points of lead was distinctly a help.

The importance of lead was demonstrated quite conclusively in the revision of a telephone directory. When such directories become unwieldy as the city grows, the necessity arises of getting more names on a page. In on instance an experiment was conducted by setting typical directory pages in different typographical arrangements and having persons look up designated telephone numbers as rapidly as possible. The arrangement finally adopted as a result involved smaller type but with 1 point lead between the lines. There were 25 per cent more names on the page and legibility increased 15 per cent. Obviously the space between the lines more than offset the decrease in the size of the type.

It is difficult to generalize on this matter of leading because it is complicated by size of type and length of line. There are indications that with the smallet type sizes leading is definitely a help. There are available specifications of limits within which it is safe to operate with reference to these three aspects of typography. One who is concerned with the more detailed specification of format would do well to consult some such recommendations.[2]

Spatial Arrangement of Page

Margins. It is seldom that every square inch of the page is used for type but it is a moot question what percentage of the area should be devoted to margins. Current practice is around 50 per cent. Conceivably we may be wasting a good bit of paper in this manner. The speed of reading method was used with material printed on a sheet with no margins or with a 22 millimeter margin. The difference in speed was less than 2 per cent and was not statistically significant. To be sure the experiment was conducted with single flat sheets and there might be a different problem in a magazine which was bound so that the pages did not lie out flat. Presumably it would be advisable to have enough margin to allow for the curvature of the paper although experiments have not been made on this particular point. There is also a possibility of distraction from things beyond the margin. One straw in the wind is an investigation of space between columns. Material was printed in two columns with about 1 millimeter or 8 millimeters between them. This difference in space be-

[2]See Paterson and Tinker, *op. cit.* p. 80.

tween the columns made no difference in speed of reading. The reader does not get outside his column very much—at least when the possible distraction is merely some more printing. It might be a different story if the adjacent page carried an interesting picture or a colored display or something that had high attention value. It is doubtful, however, if any reasonably sized margin would take care of such a factor as this. The problem would be to locate the printed material where the other distracting material with high attention value was not adjacent.

Two column arrangement. It was noted above that the reader is handicapped if the printed line is too long so that he gets lost in finding the beginning of the following line and is unable to capitalize on the premonitions of meaning in subsequent lines. The obvious remedy if the page is too wide for effective reading is to set the material in columns.

Space between columns. If material is to be set in columns there is a further problem as to how to seperate them. It is possible to leave considerable space between the columns or to put in a rule, that is a straight vertical line, or any combination of rule and space. An experiment varied the conditions from a rule with no space up to 8 millimeters space. The results were entirely negative as far as speed of reading was concerned. Hence if one puts a single rule between the columns it is unnecessary to bother with any space at all.

Paragraph arrangement. Breaking material into paragraphs promotes legibility. The speed of reading test cited above involved some 30 short paragraphs, each composed of about 30 words. These paragraphs were printed with the usual indentation. By way of variation, however, five paragraphs were thrown into one so that only every fifth one was indented. Under these circumstances legibility, was less efficient by about 7 per cent. Presumably the paragraphs corresponded somewhat to "thought units" and when the arrangement was changed the transitions were less obvious and the reader had more difficulty.

Color

On occasion it is desirable to use color in connection with printed material generally for the purpose of attracting the readers' attention. A common instance is ordinary black typography on a tinted or colored paper stock. Sometimes we find colored letters, perhaps on a colored background. There have been numerous experiments with variations of color of type and background but the principles involved seem pretty clear. It all comes down to a matter of brightness contrast between the letter and the background. It does not make so much difference whether it is a red letter on a blue ground or a blue letter on a red ground but if one is light and one is dark that is the principal consideration from the standpoint of legibility. Similarly, if black letters are to be used on colored paper, the paper should be of rather light tint in order to provide adequate contrast. Apparently "pure" white is not mandatory.

Printing Stock

A related problem deals with the surface of the paper on which the material is printed. A glossy paper makes it possible to do a better job of printing pictures by the half-tone process. However, there is the possibility that the glossy paper will

produce glare and reduce legbility. Speed of reading tests were made with the printed material on glossy white, dull coated white, or antique white. The results showed no difference in the speed of reading. Presumably if the glare is noticeable the reader changes the position of the page to minimize reflection.

Illumination

While not strictly a problem of typography, it should be pointed out in passing that the illumination of the printed page affects the legibility. One variable is the intensity, which can be changed by so simple an expedient as putting in a lamp of different wattage. Experiments on the relation between intensity of illumination and effectiveness of vision show a consistent trend with a number of methodologies. At very low intensities it is found that a comparatively small increase produces a marked increase in efficiency of vision, but at higher levels increase in intensity produces less increase in efficiency and finally a point is reached where additional intensity produces no increase in effectiveness of reading or other visual task. The significant problem is to determine the point beyond which additional intensity of illumination is unnecessary and a waste of power. Specifications have been published as to the minimum intensity recommended for various visual tasks including reading. One gets the impression from reading these various specifications or codes that much depends on who is making the recommendation. Persons who have some connection with industries that might be interested in selling electric power or equipment appear to give higher estimates than scientists who have studied the matter with purely esoteric interest. One of the former, for example, recommends 30 foot-candles for reading a newspaper whereas one of the latter states that it can be read comfortably at 7 foot candles but to play safe recommends 15 or 20 as entirely adequate. The point is that high intensities of illumination are not necessary for ordinary reading although there is a point below which one is handicapped.

Another variable is the distribution of illumination. In a direct system the fixtures are so located that practically all the light comes directly from the source to the work whereas in an indirect system, the lamp is above an opaque bowl with a glazed interior and all the light is reflected from the walls and ceiling down to the book. Numerous gradations between these extremes are possible. Experiments upon visual acuity as influenced by this type of thing have yielded, on the whole, negative results. Differences in visual performance of not over 3 per cent were found with a wide range of installations. However, there is a problem of ocular fatigue.

Experiments indicate that prolonged reading under a direct installation is much more fatiguing than under an indirect. The fatigue actually is located in the muscle which focuses the lens of the eye. What actually happens is that when there are bright sources visible in the periphery of the visual field the reader tends to fixate them and focus on them and then his attention comes back to the book and then back to the source of light. This pulling and hauling on the ciliary muscle which focuses the eye results in fatigue. There may be some too in the muscles which turn the eyeball. It can be minimized by using an installation such that the reader

is not stimulated so much by these peripheral sources. An indirect lighting system, of course, eliminates them. In other systems if the fixtures are behind him or at one side, the effect will be less pronounced. Local lighting such as a drop light can be used to produce adequate intensity on the work while the general illumination of the room can be at a much lower level and arranged to minimize the peripheral sources.

The foregoing discussion, as indicated earlier, has dealt primarily with typography from the standpoint of legibility. None of the experiments, to the writer's knowledge, have related typography uniquely to reading comprehension. The scientist is inclined to take one variable at a time and actual ability to make out the material has been the most obvious variable to investigate. The widely used speed of reading technique in which the reader has to discover the wrong word in each paragraph obviously involves some comprehension but that is rendered essentially constant by devising forms of equal difficulty. It is probable that typography's greatest effect relates to those reading habits of noting a few characteristic letters or landmarks in a word and inferring the rest from those cues.

Summary

Readability may be evaluated by photographing eye movements, by speed of reading tests, by flashing material in an exposure apparatus, by determining the maximum distance at which a thing can be read, by using a device which throws it in and out of focus and possibly by rate of blinking. With reference to type face or style, it develops that when the conditions of reading are difficult as, for example, with a distant sign, the width of the strokes and the uniformity of such width is important. For ordinary reading with the type faces in common use the effect is not so pronounced except in the case of extreme or ornate styles.

Material set in small letters is consistently more legible than the same material set in capitals largely because of the fact that we rely on the shape of the word rather than attending to every individual letter, and words in small letters have more characteristic shapes. Lines of moderate length, somewhere in the vicinity of 80 millimeters, appear most legible with the sizes of type in common use. If the line is too long, the reader gets lost in returning to the next line and is unable to utilize the premonitions of meaning which he gets in the margin of his attention because the context of two adjacent lines is too dissimilar. A little space between the lines produced by leading is often helpful, especially with smaller sizes of type. Margins are often larger than is necessary for good legibility, although if the page is somewhat curved as in certain magazines, the margin is necessary to keep the printed material from being obscured by the adjacent page. If the page is unduly wide, it is advisable to set the material in two columns. In that case a single rule between the columns appears adequate without any further space. The usual arrangement of indented paragraphs promotes legibility.

When using colored letters or colored paper or both, the important consideration is the contrast in brightness between the letters and the ground. There is little difference in readability when using a glossy paper versus an unfinished paper. The

reader evidently adapts himself to any possible reflections from the glossy surface. When planning illumination under which reading is to be done, it is advisable to avoid bright sources in the periphery of the visual field which will stimulate the reader to focus on them with resulting fatiguing conflict in the eye muscles.

Bibliography

1. Paterson, D. G. and Tinker, M. A., How to make type readable. New York: Harpers, 1940, pp. 209. A bibliography, pp. 193-201, of about 100 titles covers the important literature up to 1939. For subsequent references the following are suggested.
2. Breland, K., Legibility of newspaper headlines printed in capitals and in lower case. *J. of Applied Psychol.*, 1944, 28, 117-120.
3. English, E., A study of the readability of four newspaper headline types. *Journalism Quarterly*, 1944, 21, 217-229.
4. Halstead, W. C., A method for the quantitative reporting of eye movement. *J. of Psychol.*, 1938, 6, 177-180.
5. Luchiesh, M. and Moss, F. K., Criteria of readability. *J. of Experimental Psychol.*, 1940, 27, 256-270.
6. Paterson, D. G. and Tinker, M. A., Influence of line width on eye movements. *J. of Experimental Psychol.*, 1940, 27, 572-577.
7. Paterson, D. G. and Tinker, M. A., Influence of size of type on eye movement. *J. of Applied Psychol.*, 1942, 26, 227-230.
8. Paterson, D. G. and Tinker, M. A., Readability of newspaper headlines printed in capitals and in lower case. *J. of Applied Psychol.*, 1946, 30, 161-168.
9. Paterson, D. G. and Tinker, M. A., The effect of typography on the perceptual span in reading. *American J. of Psychol.*, 1947, 60, 388-396.
10. Paterson, D. G. and Tinker, M. A., Influence of leading upon readability of newspaper type. *J. of Applied Psychol.*, 1947, 31, 160-163.
11. Tinker, M. A. and Paterson, D. G., Readability of mixed type forms. *J. of Applied Psychol.*, 1946, 60, 631-637.
12. Uhlaner, J. E., The effects of thickness of stroke on the legibility of letters. *Proceedings Iowa Academy of Science*, 1941, 48, 319-324.

THE CONCEPT OF READABILITY

(Continued from Page 7)

9. Paterson, Donald G. and Tinker, Miles A. *How to Make Type Readable.* New York: Harper and Brothers, 1940.
10. Sterner, Alice P. *Radio, Motion Picture, and Reading Interests,* Contributions to Education No. 932. New York: Bureau of Publications, Teachers College, Columbia University, 1947.
11. Strang, Ruth. "Estimating the Difficulty of High-School and College Reading Material" (Abstract), Practical Values of Educational Research, pp. 50-51. Official Report of the American Educational Research Association, Washington, D. C.: American Educational Research Association, 1938.
12. Vogel, Mabel, and Washburne, Carleton. "An Objective Method of Determining Grade Placement of Children's Reading Material," *Elementary School Journal*, XXVIII (January, 1928), pp. 373-381.
13. Waples, Douglas, and Tyler, Ralph W. *What People Want to Read About.* Chicago: American Library Association and the University of Chicago Press, 1931.
14. Zeller, Dale. *Relative Importance of Factors of Interest in Reading Materials of Junior High School Pupils.* Contributions to Education, No. 841. New York: Bureau of Publications, Teachers College, Columbia University, 1941.
15. Bryson, Lyman. "What Are Readable Books?" *Educational Forum*, I (April 1931) pp. 408-411.
16. Tinker, Miles A., and Paterson, Donald G. "Reader Preferences and Typography," *Journal of Applied Psychology*, XXVI (January 1942) pp. 38-40.

Techniques for Selecting and Writing Readable Materials

EDGAR DALE AND JEANNE S. CHALL[1]

The previous articles in this series have considered the factors which affect the readability of printed material. This chapter shows how teachers, librarians, and community leaders can use readability research to select reading materials for specific readers. We shall also outline some techniques that writers and editors can use in writing readable materials.

The techniques which follow are based on the concept of readability proposed by the writers in their previous article:

> In the broadest sense, readability is the sum total (including the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it. The success is the extent to which they understand it, read it at an optimum speed, and find it interesting. (3, p. 23)

We are not concerned here with the important questions of content. Whether or not more emphasis should be place on colonial times in a fifth grade American history textbook, or whether high school civics texts should put more or less emphasis on youth volunteer service, are not readability problems. But the way these topics are presented so that they may be read with optimal interest, understanding, and speed is a very important problem in readability.

Some persons object to readability analyses and suggestions for writing readable materials. They say that writing is an art. We agree with them. But we also believe that art is helped by the use of related science.

Some teachers and librarians strongly favor an entirely individualized reading program. Since children within any grade have different reading abilities and interests, material that is readable for one child may not be readable for his classmate. They suggest, therefore, that the problem of readability can be solved by supplying a wide variety of materials so that every child can select what is "readable for him."

However, even if we were to accept completely the philosophy of "to each his own criterion of readability," someone must still select the history text, the basic reader, the geography book, etc. And even when a wide variety of materials is available for each class, these materials must be selected to meet the abilities and interests of the children in the class.

How are such selections to be made so that they can be read? What objective standards should the teacher, librarian, or group leader use to select or recommend books that will fit the abilities and interests of the group? How should a 6th grade science book be written so that sixth graders will profit from it? Here the re-

[1]Members of the faculty of the School of Education, Ohio State University.

search findings in readability can help those who select, recommend, or write reading materials.

One more caution: the last word has not yet been said about readability. We have objectively analyzed only a few of the factors influencing readability. The other factors should be considered, if only for the purpose of subjective evaluation.

In selecting and writing readable materials we consider three questions:

1. Who will read the material?
2. What is the purpose of the material?
3. Does the material fit the readers and the purpose?

The first question seems elementary, but it is not so simple. If it were, we would not be able to cite such illustrations as the following. This is the concluding paragraph of a booklet on cancer written specifically for children in grades 5 and 6:

> Take pains to prevent *chronic irritations* to any part of the body. This should include protection from over-exposure to sun and wind, repair of jagged teeth, avoiding *ill-fitting dental plates, temperance in the use of tobacco, correction of chronic constipation.* In other words, avoid unnecessary abuse of any *tissues.* (italics ours)

Even if 5th and 6th graders knew what *temperance, chronic irritations,* and *chronic constipation* meant, of what interest value is it to give them information about *ill-fitting dental plates?* Such examples can be multiplied.

The important thing to remember is that answering the simple question of "Who will read the material?" involves a knowledge of many characteristics of the prospective readers. These characteristics include: reading abilities, interests, age, sex, intellectual maturity, background of information, etc.

The larger the group of readers, the harder it is to get a clear picture of these characteristics.

It is harder, too, to determine the characteristics of readers as their age increases. It is harder to get a picture of the *who* when we select or write for high school pupils than for elementary pupils. It is harder still for members of an adult education class, and still harder when we write or select materials for the "general adult" reader. Differences in reading ability, experience, interest, and intellectual maturity tend to become greater as the average age of the group increases.

Defining the Audience

Teachers, librarians, community leaders, and authors can use the techniques suggested in the following outline to help them judge some important characteristics of the readers for whom materials are being selected or written.

1. How well does the group read?

Children:

1. Give a standardized reading test to the group. Change the scores into reading grades. Note the wide variation in grade scores.
2. Compute the average reading grade for the class. Also note those children at the upper and lower limits so that special materials of appropriate difficulty can be provided for them.
3. Use informal reading tests such as word-recognition tests and observe children's performance on books of known reading difficulty.

Adults: 1: When a reading test can be given, as in an adult education class, obtain the average and range of scores.

2. When a reading test cannot be given, use the last grade of formal schooling as an estimate of reading ability. Average these for the group and note those at the extremes who may need special materials.

3. When selecting or writing materials for a large heterogeneous group such as parents of children in school, or for the "average adult" in a community, consult the census data on the " Percentage of Persons 25 years old and over by years of school completed," in the 1940 Population, Vol. II, *Characteristics of the Population*. Part 1 contains the statistics for the U. S. as a whole, and Parts 2 to 7 give the statistics for states, cities, and counties. These are further broken down by sex, race, native and foreign born, urban and rural.

2. What are the reading interests, habits, and tastes of the group?

Children: 1: In general, be familiar with the major changes in reading interests as children proceed through the primary, middle grades, and high school level. See Witty (33) for a review of important studies.

2. More specifically, be familiar with the findings of the studies on reading interests of the major age groups. The literature is voluminous but the following may be a good start:

Elementary school:
General reading interests—Witty (33) for a review of findings of important studies. See also Gates (11), Lazar (15), Thorndike (28), Rankin (26).
Reading interests in special fields—Von Qualen (29) and Williams (32) for children's interest in science, Kyte (13) for children's interests in poetry.

High School:
General reading interests—Witty (33), LaBrant and Heller (14), Zeller (35), Sterner (27). Reading interests in special fields—47th NSSE Yearbook (23), Gray (12).

3. Judge the interests of your own group through informal discussions, questionnaries, etc.

Adults: 1. Consult studies on adult interests. A study of Waples and Tyler (30) is a good way to begin. Besides giving some major interests of specific adult groups, they propose a scheme for discovering reading interests. See also recent studies of adult reading habits—Link and Hopf (16), and the National Opinion Research Center's Study for the American Library Association (22).

2. In addition, make inquiries of your own to find out what the group does with its spare time, what books and magazines they like to read, etc. Draw up a simple questionnaire or use informal discussions.

3. How much does the group already know about what they are to read?

This question tries to get at the information or misinformation that readers al-

ready have on a particular subject. Some estimate of their present information will help determine whether the material under consideration is too elementary or too advanced. Readers' knowledge of a subject usually indicates their interest in a subject That is, the more the readers already know about the subject, the more interested they will probably be in it. This question is particularly important in selecting and writing expository material.

Children:
1. At the lowest grades, we usually assume a minimum of knowledge on most subjects and start from "scratch," building up important concepts. However, even here there is a tendency to introduce unkown abstractions. See Ordan and Lorge (25) on children's social concepts.
2. If possible, give a short test or hold discussions to find out whether the children understand some of the basic concepts upon which the book is based. Record the results. When in doubt, assume that the information is not known.

Adults:
1. Information surveys run by public opinion polling groups provide valuable hints of adults' information and misinformation on many social issues. (See *The Public Opinion Quarterly's* regular reviews of all available results of polls in the United States.)
2. Information on most subjects is positively related to socio-economic level, occupational status, and educational achievement. Thus, if we are writing a pamphlet on the UN for unskilled workers, we would assume that they know less about the UN than a group of foremen.
3. If material is for an adult education class, a formal test or an informal discussion can be held to determine their knowledge of necessary basic concepts. There is always a danger that we will overestimate the readers' information and underestimate their intelligence.

The above breakdowns for defining the audience are not complete. They help the teacher, librarian, group leader, and writer become aware of the variety of characteristics to be considered in describing an audience to provide them with suitable materials.

Defining the Purpose

The question, "What is the purpose of the material?" must also be answered before any selecting or writing is begun.

Here we must face such questions as these:

1. Is the material to serve as a broad orientation to the field or is it to be studied in detail?
2. Is the purpose of the material to communicate definite steps to action or is it to change attitudes, supply information, reach appreciation? Or is it designed to do a combination of these?
3. Is the book to be read without help from the teacher?
4. Is the reading material to be required or merely recommended?
5. Is the book to be used as a text for the entire class? Will other books or materials be available?

Of course, the purpose cannot be defined by the teacher alone. It depends upon the broad and specific objectives of the curriculum, the financial condition of the particular school, and the availability of materials on particular subjects. It also depends upon the reader's specific purpose in reading. He may be reading for entertainment, for getting answers to specific ques-

tions, for solving problems, for specific directions for doing something, etc. However, answers to the above questions in terms of readability provide some standards against which materials can be evaluated.

For example, a book to be read without the teacher's help, should be easier than one which is to be studied in class. When only one textbook is used in geography, that book should be within the comprehension of *most* of the class. We use the word *most* rather than *all* because there may well be a few extremely poor readers in the class who might have difficulty even with the simplest book. Books that are to be read for general import and appreciation can be harder than those read for detailed information.

Analyzing the Material

Once we have answered the questions of "Who is going to read the material?" and "What is the purpose?", we can proceed to the third question, "Does the material fit the readers and the purpose?"

At this stage, however, it is best to treat the selection and writing of materials separately. First, we shall take the selection of materials already written.

1. How difficult is the material?

A readability formula can be used to get a rough estimate of the difficulty of a book, pamphlet or article. However, we must realize that the available formulas measure only one aspect of difficulty—expressional or structural difficulty. Only such factors as vocabulary and sentence structure are measured. The readability formulas do not directly measure conceptual difficulty, organization of the material, abstractness of subject matter—all known to affect comprehensibility. Results from formulas should therefore be interpreted cautiously.

For a more comprehensive discussion of the uses and limitations of readability formulas, see the second article in this series by Dr. Lorge (19).

The following outline suggests the formulas that will be most useful for estimating the difficulty of different kinds of materials.

Children:
1. For estimating the difficulty of children's voluntary reading the Washburne-Morphett formula can be used (31). In the *Right Book for the Right Child* over 1000 books known to be liked by children and having literary value have already been graded in difficulty (1).
2. The Lorge formula can be used to estimate the comprehension difficulty of the kind of reading pupils do in school (18).
3. To estimate the difficulty of texts planned for children below the fourth or fifth grade, Lorge (19) suggests a weighted index of vocabulary load. See Yoakum (17), Dolch (6 and 7). See MacLatchy (20) for estimating difficulty of preprimers.

Adults:
1. Flesch (9) and Dale-Chall (13) formulas will give adequate estimates of difficulty of most materials. For specific subject matter areas, see Ojemann (24) for estimating difficulty of parent education materials and Dale-Chall for health materials.
2. Books for adult beginners have recently been graded in difficulty by Fihe and others (10).

Note: Book lists for the different grades and age groups are published regularly. In addition, publishers give estimates of the grade placements of their books. Before these estimates are accepted, make certain that you know on what basis the books have been "graded."

2. Is the material suitable in difficulty for the readers?

An answer to this question depends not only upon the predicted difficulty of the material and the abilities of the readers, but also upon the purpose served by the material.

Suppose we are selecting a reading text for a 5th grade class. The reading abilities, according to an adequate reading test, may range from 3rd grade through the 8th grade, with an average of about 5. The textbook being considered has a readability index by the Lorge formula of 5.2. Is this suitable for the class in terms of comprehensibility? The answer is yes and no. Yes, if we are concerned with only the upper half of the class, who will be able to read it with ease and understanding. No, because it may be too hard for the lower half.

If this is the only text to be used, we must decide whether it should not actually be easier than 5.2. Perhaps 4.5 or 4.0 would be better so that many more children in the class will have success with it.

If we do select a text that is a half or full grade below the average reading ability of the group, will this cut off the better readers? Will they find the book unchallenging? Will it retard their progress in reading? We believe not. The better readers can and do read material below their grade level with interest and profit. The teacher can assure their progress in reading by making available or recommending other more difficult material.

The children who read several grades below the class average are a greater problem. Even if the basic text is a half or full grade below the class reading level, it will be too difficult for them. The teacher must provide them with other materials that they can read without frustration.

In the subject matter areas, selecting a book that will be within the comprehension of the lower half of the class is even more important. Since our chief purpose is to impart information, rather than to give practice in reading, we must make certain that no unnecessary difficulties keep the less able readers from learning their history, geography, and science.

In selecting materials for adults, we have a similar problem. What level of difficulty will fit most readers and yet be challenging enough for the more able? If there is a choice between two books that cover the same subject, we would select the easier version if it is well written. It will not only be appreciated by the less able readers but will not be resented by the abler readers.

We found this to be true in interviewing readers' reactions to the National Tuberculosis Association booklet *Your Baby*. This booklet was at the 7-8th grade level by the Dale-Chall formula. We found that mothers who had graduated from college liked it as much as those who had not gone beyond the 7th grade. This indicates that material does not have to be at the grade level of the particular readers. In fact, even

for children, we must be flexible in our interpretation of the formula grade placements. Because we find a book readable at the 5th or 6th grade level, it does not mean that 8th graders will not enjoy it and profit from it. Actually interpreting readability scores as grade levels may be misleading. A grade level of 7 or 8 may merely mean that the material is clear, concrete, and skillfully written—not that it is child-like in its subject-matter and approach.

Thus far, we have discussed two questions regarding reading materials: "How difficult is the material?" and "Is the material suitable in difficulty for the readers?" We shall now discuss the questions of interest, organization, and typography.

3. Will the material be interesting to the readers?

For children in the primary grades, the interest value of the subject matter and the treatment must be studied. Books should be evaluated in the light of the research findings on children's reading interests. The specific interests of the readers for whom the selection is being made must also be considered.

For subject matter materials, the curriculum determines the content. We do not usually ask whether children in the 4th grade will be interested in reading about Columbus or about Colonial times. However, we can ask whether they will find the treatment of this particular information interesting.

At any rate, we should estimate readers' interests in both subject matter and treatment. If we have to say, "They really aren't interested in the agricultural products of Australia," then the treatment will have to carry the burden of creating interest.

An oversimplified rule is: the more the geography book relates Australian agriculture and industry to the children's lives, the more concrete the treatment, the more personalized the approach, the more interesting they will find it.

4. Is the material well organized?

We do not yet have objective tools by which to answer this question but the reader can estimate it. In a sense good organization implies that one idea leads into the next. There is a building up of sequence, then the tying together in the conclusion. Well-organized material, whether narrative or expository, has a good flow and ordering of ideas.

5. Is the typography adequate for optimum speed of reading and interest?

Are the visual illustrations (pictures, charts, etc.) relevant? Are they adequately captioned and sufficiently explained in the textual material? Is the general appearance of the book or pamphlet attractive and suitable for the content and treatment?

For judging the adequacy of typography, see the previous article by Dr. Burtt (2). Recently Malter (21) reviewed significant findings on children's preferences of drawings and pictures.

Writing Readable Materials

The problem of writing readable materials varies so much with the subject matter written about, the age or grade written for, and the purpose of the material, that we shall present here only some general suggestions for attacking the problem.

First, the writer must know for whom he is writing. He must also have a clear idea of what he is trying to get across. Then, if he can picture one or two typical readers, and tell his "message" to them, he will have a better chance of making himself understood and interesting.

The suggestions by Flesch in the *Art of Plain Talk* (8) and by Dale and Hager in their recent article, "How to Write to Be understood" (5) should be helpful to writers. The article by Dale and Hager is particularly applicable to the writing of technical materials.

In preparing the first draft, the writer should not be distracted by word lists, sentence length, prepositional phrases. He should tell his story," keeping in mind the need to be understood by the majority of the audience for whom he is writing. After the first draft, he can evaluate what he has written by using some of the techniques outlined above in *Analyzing the Material*. He can then revise in light of the interests, abilities, and intellectual maturity of his prospective readers.

Summary

Under the three questions: "Who will read the material?" "What is the purpose?" and "Does the material fit the readers and the purpose?" we have suggested some techniques for selecting and writing readable materials.

Before a teacher, librarian, or group leader analyzes material to find whether it is readable, he must have a clear picture of the characteristics of his readers and the purpose which the material will serve. An author who wishes to write readable materials, must also know his audience and the purpose of the material. We have therefore suggested some procedures which the teacher or writer can use to define his audience and purpose. We have also referred the reader to publications which will help him analyze materials to find whether they fit his reader's abilities, interests, and intellectual maturity.

Bibliography

1. American Library Association and the Winnetka Public Schools. *The Right Book for the Right Child*. New York: The John Day Company, 1942 (3rd edition).
2. Burtt, Harold E. "Typography and Readability," *Elementary English*, XXVI (April, 1949), pp. 212
3. Dale, Edgar, and Chall, Jeanne S. "The Concept of Readability," *Elementary English*, XXVI (January, 1949), pp. 19-26.
4. Dale, Edgar, and Chall, Jeanne S. "A Formula for Predicting Readability," and "Instructions," *Educational Research Bulletin*, XXVII (January 21, 1948 and February 17, 1948), pp. 11-20 and 37-54.
5. Dale, Edgar, and Hager, Hilda. "How to Write to Be Understood," *Educational Research Bulletin*, XXVII (November 10, 1948), pp. 207-216.
6. Dolch, E. W. *Problems in Reading*. Champaign, Ill.: Garrard Press, 1948, Chapter XXI.
7. Dolch, E. W. "The Use of Vocabulary Lists in Predicting Readability and in Developing Reading Materials," *Elementary English*, XXVI (March, 1949), pp. 142.
8. Flesch, Rudolf. *The Art of Plain Talk*. New York: Harper and Bros., 1946.
9. Flesch, Rudolf. "A New Readability Yardstick," *Journal of Applied Psychology*, XXXII (June, 1948), pp. 221-233.
10. Fihe, Pauline J., Wallace, Viola, and Schulz, Martha, compilers. *Books for Adult Beginners; Grades I to VII*. Rev. Ed. Chicago: American Library Association, 1946.
11. Gates, Arthur I. *Interest and Ability in Reading*. New York: The MacMillan Company, 1930.
12. Gray, William S., Ed., *Improving Reading in Content Fields*, Supplementary Educational Monographs, Illinois: The University of Chicago Press, 1947.

13. Kyte, G. C. "Children's Reactions to Fifty Selected Poems," *Elementary School Journal*, XLVII (February, 1947), pp. 331-339.
14. LaBrant, Lou L., and Heller, Frieda M. *An Evaluation of Free Reading in Grades 7 to 12, Inclusive*. Contributions to Education, No. 4. Columbus: Ohio State University, 1939.
15. Lazar, May. *Reading Interests, Activities, and Opportunities of Bright, Average and Dull Children*. Teachers College Contributions to Education, No. 707, New York: Bureau of Publications, Teachers College, Columbia University, 1937.
16. Link, Henry C. and Hopf, Harry A. *People and Books: A Study of Reading and Book-Buying Habits*. New York: Book Industry Committee, Book Manufacturers' Institute, 1946.
17. Lorge, Irving. "Predicting Reading Difficulty of Selections for Children." *Elementary English Review*. XVI (October, 1939), pp. 229-233.
18. Lorge, Irving. "The Lorge and Flesch Readability Formulae: A Correction," *School and Society*, LXVII (February 21, 1948), pp. 141-142. See also "Predicting Readability," *Teachers College Record*, XLV (March, 1944), 404-419.
19. Lorge, Irving. "Readability Formulae—An Evaluation," *Elementary English*, XXVI (February, 1949), pp. 86-95.
20. MacLatchy, Josephine H. and Wardwell, Frances. "Common Pre-Primer Words," *Educational Research Bulletin*, XXVII (November 10, 1948), pp. 199-206.
21. Malter, Morton S. "Children's Preferences for Illustrative Materials," *Journal of Educational Research*, XLI (January, 1948), pp. 378-385.
22. National Opinion Research Center. *What, Where, and Why Do People Read*. Report No. 28, January, 1946.
23. National Society for the Study of Education, *Forty-Seventh Yearbook*, Part II. *Reading in the High School and College*, Chicago: The University of Chicago Press, 1948. See especially chapters VII, VIII, and IX.
24. Ojemann, Ralph H. "The Reading Difficulty of Parent Education Materials," *Proceedings of the Iowa Academy of Sciences*, XL, (April, 1933), pp. 159-170. See also "The Reading Ability of Parents and Factors Associated with Reading Difficulty of Parent Education Materials," Researches in Parent Education II, *University of Iowa Studies, Studies in Child Welfare*, VIII (March, 1934), pp. 11-32.
25. Ordan, Harry and Lorge, Irving, *Social Concepts and the Child Mind*. New York: Columbia University Press, 1945.
26. Rankin, Marie. *Children's Interests in Library Books of Fiction*. Teachers College Contributions to Education, No. 906. New York: Teachers College, Columbia University, 1944.
27. Sterner, Alice P. *Radio, Motion Picture, and Reading Interests: A Study of High School Pupils*. Teachers College Contributions to Education, No. 932, New York: Bureau of Publications, Teachers College, Columbia University, 1947.
28. Thorndike, Robert L. *Children's Reading Interests*, New York: Teachers College, Columbia University, 1941.
29. Von Qualen, V. D. and Kambly P. E., "Children's Interest in Science as Indicated by Choices of Reading Materials," *School, Science, and Mathematics*, XLV (December, 1945), pp. 798-806.
30. Waples, Douglas, and Tyler, Ralph W. *What People Want to Read About*. Chicago: American Library Association and the University of Chicago Press, 1931.
31. Washburne, Carleton, and Morphett, Mabel Vogel. "Grade Placement of Children's Books," *Elementary School Journal*, XXXVII (January, 1938) pp. 335-364.
32. Williams, Alice M. *Children's Choices in Science Books*. Child Development Monograph, No. 27. New York: Teachers College, Columbia University, 1939.
33. Witty, Paul. *Reading in Modern Education*, Boston: D. C. Heath and Company, 1949, Chapters II and V.
34. Yoakam, G. A. *Revised Directions for Using the Yoakum Readability Formula*, Unpublished Manuscript, University of Pittsburgh, March, 1948.
35. Zeller, Dale. *Relative Importance of Factors of Interest in Reading Materials of Junior High School Pupils*. Contributions to Education, No. 841, New York: Bureau of Publications, Teachers College, Columbia University, 1941.